

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:34:53 ; Search time 47 Seconds  
(without alignments)  
1983.752 Million cell updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNEPEQHRSMPKPKKNA.....LPAAVTWLILSVISSYPS 297

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1563	100.0	297	12	US-09-853-641-2		Sequence 2, Appli
2	1495	95.6	285	9	US-09-738-626-5716		Sequence 5716, Ap
3	1495	95.6	285	12	US-10-627-476-100		Sequence 100, App
4	1064	68.1	292	12	US-10-282-122A-53970		Sequence 53970, A
5	725.5	46.4	311	12	US-10-282-122A-62176		Sequence 62176, A
6	719	46.0	306	12	US-10-282-122A-62583		Sequence 62583, A
7	719	46.0	306	12	US-10-282-122A-64789		Sequence 64789, A
8	714	45.7	312	12	US-10-282-122A-63884		Sequence 63884, A
9	551	35.3	392	14	US-10-156-761-10160		Sequence 10160, A
10	310	19.8	271	12	US-10-282-122A-69792		Sequence 69792, A
11	308	19.7	285	12	US-10-282-122A-55991		Sequence 55991, A
12	306	19.6	79	11	US-09-824-408A-906		Sequence 906, App
13	303	19.4	273	12	US-10-282-122A-49018		Sequence 49018, A
14	299	19.1	271	12	US-10-282-122A-67557		Sequence 67557, A
15	299	19.1	282	12	US-10-282-122A-77963		Sequence 77963, A

16	298	19.1	280	12	US-10-282-122A-77418		Sequence 77418, A
17	298	19.1	285	12	US-10-282-122A-75280		Sequence 75280, A
18	298	19.1	285	12	US-10-282-122A-76157		Sequence 76157, A
19	295	18.9	263	12	US-10-282-122A-59487		Sequence 59487, A
20	292.5	18.7	248	12	US-10-282-122A-73005		Sequence 73005, A
21	288.5	18.5	265	12	US-10-282-122A-65186		Sequence 65186, A
22	286.5	18.3	265	12	US-10-282-122A-65594		Sequence 65594, A
23	285	18.2	249	12	US-10-282-122A-56426		Sequence 56426, A
24	285	18.2	249	13	US-10-136-517-17		Sequence 17, Appl
25	283	18.1	265	12	US-10-282-122A-61125		Sequence 61125, A
26	282.5	18.1	287	12	US-10-282-122A-68659		Sequence 68659, A
27	280.5	17.9	289	12	US-10-282-122A-67464		Sequence 67464, A
28	277	17.7	288	12	US-10-282-122A-58351		Sequence 58351, A
29	276	17.7	273	12	US-10-282-122A-47625		Sequence 47625, A
30	276	17.7	273	12	US-10-282-122A-50187		Sequence 50187, A
31	275	16.4	274	12	US-10-282-122A-44770		Sequence 44770, A
32	249.5	16.0	260	12	US-10-282-122A-71462		Sequence 71462, A
33	249	15.9	241	12	US-10-282-122A-54574		Sequence 54574, A
34	248	15.9	282	12	US-10-282-122A-63044		Sequence 63044, A
35	247.5	15.8	260	12	US-10-282-122A-70940		Sequence 70940, A
36	247	15.8	293	15	US-10-369-493-16004		Sequence 16004, A
37	247	15.8	309	15	US-10-369-493-16356		Sequence 16356, A
38	247	15.8	309	15	US-10-369-493-15629		Sequence 15629, A
39	246.5	15.8	311	9	US-09-815-242-11873		Sequence 11873, A
40	246.5	15.8	311	12	US-10-282-122A-66387		Sequence 66387, A
41	244.5	15.6	264	12	US-10-282-122A-72396		Sequence 72396, A
42	243	15.5	285	12	US-10-282-122A-51478		Sequence 51478, A
43	239.5	15.3	266	12	US-10-282-122A-58551		Sequence 58551, A
44	237.5	15.2	260	12	US-10-282-122A-48529		Sequence 48529, A
45	237.5	15.2	296	15	US-10-369-493-19811		Sequence 19811, A

ALIGNMENTS

RESULT 1  
US-09-853-641-2  
; Sequence 2, Application US/09853641  
; Publication No. US20040092710A1  
; GENERAL INFORMATION:  
; APPLICANT: NAMPOOTHIRI, Madhavan  
; TITLE OF INVENTION: Nucleotide Sequences Coding for the CdsA Gene  
; FILE REFERENCE: 032301 WD 1171  
; CURRENT APPLICATION NUMBER: US/09/853,641  
; CURRENT FILING DATE: 2001-08-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-853-641-2

Query Match 100.0%; Score 1563; DB 12; Length 297;  
Best Local Similarity 100.0%; Pred. No. 3.6e-144;  
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNEPEQHRSMPKPKKNAAGDLKAAIAGVGLGVLLGLVLSFWGWIYLVAGFMAAA	60
Db	1	MNEPEQHRSMPKPKKNAAGDLKAAIAGVGLGVLLGLVLSFWGWIYLVAGFMAAA	60
Qy	61	TWEVGSRLKEGGYHLLPLPIMIIGGQAIWLSPFFGTMGILASFVATVLVLMYFRIFYNGT	120
Db	61	TWEVGSRLKEGGYHLLPLPIMIIGGQAIWLSPFFGTMGILASFVATVLVLMYFRIFYNGT	120
Qy	121	EKEARNYLDTSVGLFVLTVLPLFGSFAALMSLMQNNSTPGTYFILTFLMCLVIASDVGY	180
Db	121	EKEARNYLDTSVGLFVLTVLPLFGSFAALMSLMQNNSTPGTYFILTFLMCLVIASDVGY	180
Qy	181	IAGVFFGSHPMAPLVSPKXSWEGFAGSIVLGSVTGALSVEHLLDHHWNGVILGALVVC	240
Db	181	IAGVFFGSHPMAPLVSPKXSWEGFAGSIVLGSVTGALSVEHLLDHHWNGVILGALVVC	240

QY 241 ATGLDVSQKRDIGIKDMSNLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 297  
 Db 241 ATGLDVSQKRDIGIKDMSNLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 297

## RESULT 2

US-09-738-626-5716  
 ; Sequence 5716, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENO, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OKAZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: Patent in ver. 3.0  
 ; SEQ ID NO 5716  
 ; LENGTH: 285  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-5716

Query Match 95.6%; Score 1495; DB 9; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 MPKPKNAGRDLCAAIAGVIGLGVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 72  
 Db 1 MPKPKNAGRDLCAAIAGVIGLGVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 60

QY 73 YHLPPIIMIGQAIILWSWPGFTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 132  
 Db 61 YHLPPIIMIGQAIILWSWPGFTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 120

QY 133 VGIFVLTWIPFGSFAAMLSLQNNISIPGTFFILTFMLCVIASDVGGYIAGVFFGSHPMA 192  
 Db 121 VGIFVLTWIPFGSFAAMLSLQNNISIPGTFFILTFMLCVIASDVGGYIAGVFFGSHPMA 180

QY 193 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWNGVILGCALVVCATLGLDLSVESQFK 252  
 Db 181 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWNGVILGCALVVCATLGLDLSVESQFK 240

QY 253 RDLGKDMNSLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 297  
 Db 241 RDLGKDMNSLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 285

## RESULT 3

US-10-627-476-100  
 ; Sequence 100, Application US/10627476  
 ; Publication No. US20040030116A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pompeius, Mark  
 ; APPLICANT: Kroger, Burkhard  
 ; APPLICANT: Schoder, Hartwig  
 ; APPLICANT: Zelder, Oskar  
 ; APPLICANT: Haberer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
 ; FILE REFERENCE: BGI-125PCPN  
 ; CURRENT APPLICATION NUMBER: US/10/627,476  
 ; CURRENT FILING DATE: 2003-07-25  
 ; PRIOR APPLICATION NUMBER: 09/602,787  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: USSN 60/141031  
 ; PRIOR FILING DATE: 1999-06-25  
 ; PRIOR APPLICATION NUMBER: DE 19931454.3  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19931478.0  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19931563.9  
 ; PRIOR FILING DATE: 1999-07-08  
 ; PRIOR APPLICATION NUMBER: DE 19932122.1  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932180.9  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932124.8  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932125.6  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932128.0  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: DE 19932180.9  
 ; PRIOR FILING DATE: 1999-07-09  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 678  
 ; SEQ ID NO 100  
 ; LENGTH: 285  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-10-627-476-100

Query Match 95.6%; Score 1495; DB 12; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 MPKPKNAGRDLCAAIAGVIGLGVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 72  
 Db 1 MPKPKNAGRDLCAAIAGVIGLGVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 60

QY 73 YHLPPIIMIGQAIILWSWPGFTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 132  
 Db 61 YHLPPIIMIGQAIILWSWPGFTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 120

QY 133 VGIFVLTWIPFGSFAAMLSLQNNISIPGTFFILTFMLCVIASDVGGYIAGVFFGSHPMA 192  
 Db 121 VGIFVLTWIPFGSFAAMLSLQNNISIPGTFFILTFMLCVIASDVGGYIAGVFFGSHPMA 180

QY 193 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWNGVILGCALVVCATLGLDLSVESQFK 252  
 Db 181 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWNGVILGCALVVCATLGLDLSVESQFK 240

QY 253 RDLGKDMNSLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 297  
 Db 241 RDLGKDMNSLLPGHGLMDRLDGLMPLAAWVTLILSVLSYSPS 285

## RESULT 4

US-10-282-122A-53970  
 ; Sequence 53970, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53970
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53970

Query Match      68.1%; Score 1064; DB 12; Length 292;
Best Local Similarity 70.5%; Pred. No. 1.9e-95;
Matches 198; Conservative 33; Mismatches 50; Indels 0; Gaps 0;

Qy 13 MKPKNNAGRDLDKAAIYAGIGLVILGIVLSPGWLIVAGFMAAATWVGSRLEGG 72
Db 11 LPKKNAGRLKAAISVIGLIGLVILGIVLSPGWLIVAGFMAAATWVGSRLEGG 70

Qy 73 YHLPFIMIGQQAIIWLSWPFMTGILASFAVATVILMYRIFRYNGTEKEARNYLRDTS 132
Db 71 YLLQSRWMLIGQVNLWLSWPFMTGILASFAVATVILMYRIFRYNGTEKEARNYLRDTS 130

Qy 133 VGIFVLTWIPFGSFAAMLSLMQNNISISCTYFILTFLMCLVIASDVGGYIAGVFGSHPMA 192
Db 131 VAIFVLTWIPFGSFAAMLSLUFETAFGKYFIVTFMLCVIASDVGGYIAGVFGSHPMA 190

Qy 193 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDDHHWMMGVILGCALVVCATLGLDLVESQFK 252
Db 191 PAVSPKSWEGFAGSIVLGSVTGALSVEHFLDDHHWMMGVILGCALVVCATLGLDLVESQFK 250

Qy 253 RDLGKQNSNLLPHGGMLMDRLDGLMPLPAAMTWLILSVISS 293
Db 251 RELGKQNSAIIPLGCGMLMDRLDGLMPLPAAMTWLILSVISS 291

RESULT 5
US-10-282-122A-62176
; Sequence 62176, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

```

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62176
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62176

Query Match      46.4%; Score 725.5; DB 12; Length 311;
Best Local Similarity 47.5%; Pred. No. 2.4e-62;
Matches 145; Conservative 47; Mismatches 94; Indels 19; Gaps 4;

Qy 2 NEPE-----QHRSRMKPKKNAGRDLDKAAIYAGIGLVILGIVLSPGWLIVYA 54
Db 10 DEPEHAVENTTEGAAGRAKKTSRAGRDLDRAAIAVGAGIAGVILVLPAPFVPIVA 69

Qy 55 GFMAAATWVGSRLEKGGYHLPIMIGQAIIWLSWPFMTGILASFAVATVILMYR 114
Db 70 MALIVASHEVVRLEAGYVIVPILLAGGQTLVLTWTFPHAGALAGFGVTVVACLFWR 129

Qy 115 IFYNGTEKE-----ARNYLRDTSVGIFVLTWIPFGSFAAMLSLMQNNISIPGTYFIL 166
Db 130 LFQMDNRKRPEPPAGSPSANYLRDASATVFLACWVLPFASFAALLVYPADGA--GRVFCL 187

Qy 167 TFMCLVIASDVGGYIAGVFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVEHFLDDHH 226
Db 188 --NITVVASDVGGYIAGVFLFGKHPMPALSPKSWEGLAGSLVLTGTAATLAATFLAGKA 245

Qy 227 WMMGVILGCALVVCATLGLDLVESQFKRDLGIKDMSNLLPHGGMLMDRLDGLMPLPAAMTWL 286
Db 246 PWVGALLGVVLTCTGLDLVESQVKRDLGIKDMSNLLPHGGMLMDRLDGLVPSAAVAWT 305

Qy 287 ILSVI 291
Db 306 VLTIV 310

RESULT 6
US-10-282-122A-62583
; Sequence 62583, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

```

```

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62583
; TYPE: PRT
; LENGTH: 306
; ORGANISM: Mycobacterium bovis
; US-10-282-122A-62583

Query Match 46.0%; Score 719; DB 12; Length 306;
Best Local Similarity 47.7%; Pred. No. 1e-61;
Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY 2 NEPEQHRSRM--PKPKNNAGRDLEKAAIAVGIGLVLLGLVLSPMGWYILVAGFMAA 59
DB 10 NPAEQPARGAKQOPATETSRAGRDLEKAAIAVGIGLVLLGLVLSPMGWYILVAGFMAA 69
QY 60 ATWEVGSRLKEGGYHLPLPIMIIGQAAIWLSPFGTNGILASFVATVLMYPRIFY-- 117
DB 70 ATHEVVRRLREAGYILPVIPLLIQGAQAVLWTPFGAVGALAGFGGVMVVCMIWRLEFMD 129
QY 118 -----NGTEKEARNYLRDTSVGIPLTWIPLFGSFAAMLQNNNSIPGTYFILTMLC 171
DB 130 SVTRPTTGAPSPGNLSDVSATVFLAVVPLFCFSGAMLVYPEN----GSGWVFCMMIA 185
QY 172 VIASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 231
DB 186 VIASDVGGYAVGVLFEGKHPVPTISPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 245
QY 232 ILGCALVVCATIGDIVESQFKDGLGKDNMNLPGHGLMDRLDGLMPLPAAMVTLISVI 291
DB 246 LLGVLFVLTALGLDIVESQVKRDGLGKDNMNLPGHGLMDRLDGLMPLPAAMVTLISVI 305

RESULT 7
US-10-282-122A-64789
; Sequence 64789, Application US/10282122A
; Publication No. US20040029129A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 64789
; TYPE: PRT
; LENGTH: 306
; ORGANISM: Mycobacterium tuberculosis
; US-10-282-122A-64789

Query Match 46.0%; Score 719; DB 12; Length 306;
Best Local Similarity 47.7%; Pred. No. 1e-61;
Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY 2 NEPEQHRSRM--PKPKNNAGRDLEKAAIAVGIGLVLLGLVLSPMGWYILVAGFMAA 59
DB 10 NPAEQPARGAKQOPATETSRAGRDLEKAAIAVGIGLVLLGLVLSPMGWYILVAGFMAA 69
QY 60 ATWEVGSRLKEGGYHLPLPIMIIGQAAIWLSPFGTNGILASFVATVLMYPRIFY-- 117
DB 70 ATHEVVRRLREAGYILPVIPLLIQGAQAVLWTPFGAVGALAGFGGVMVVCMIWRLEFMD 129
QY 118 -----NGTEKEARNYLRDTSVGIPLTWIPLFGSFAAMLQNNNSIPGTYFILTMLC 171
DB 130 SVTRPTTGAPSPGNLSDVSATVFLAVVPLFCFSGAMLVYPEN----GSGWVFCMMIA 185
QY 172 VIASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 231
DB 186 VIASDVGGYAVGVLFEGKHPVPTISPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 245
QY 232 ILGCALVVCATIGDIVESQFKDGLGKDNMNLPGHGLMDRLDGLMPLPAAMVTLISVI 291
DB 246 LLGVLFVLTALGLDIVESQVKRDGLGKDNMNLPGHGLMDRLDGLMPLPAAMVTLISVI 305

RESULT 8
US-10-282-122A-63884

```



```
; Sequence 63884, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63884
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-282-122A-63884
```

```
Query Match 45.7%; Score 714; DB 12; Length 312;
Best Local Similarity 49.5%; Pred. No. 3.2e-61;
Matches 146; Conservative 49; Mismatches 84; Indels 16; Gaps 5;

QY 9 RSMRM-PPKNNAGRDLEKAAIANGVIGLVLLGVLSPNGWYILVAGFMAAATWEVGSR 67
Db 21 RAMROSTKNTTPRAGNLPAAIANGVIGLVLLGVLSPNGWYILVAGFMAAATWEVGSR 80
QY 68 LKEGGYHLPLPIMTIIGQAIWLSPFGTGMILASVFATVILVMYFRIFY--NTEKEAR 125
Db 81 LREAGYVIPALPLLGGQFTWLTWPYTVGALAGFGATVVCWMLRVLVHMDNSKQHSR 140
QY 126 -----NVLRTSVGIFVLTWIPFGSFAAMLSMNNNSIPGTFFILTMCLVIASD 176
Db 141 EALAGPPVSNLRTASATVFLAAWVLPFASFAALLVVPKDA--GRVFCFL--MIAVVASD 196
QY 177 VGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVVTGALSVEHFLDDHWMGVILGCA 236
Db 197 VGGYTVGVLFKXHLVPRISNKSWEFGAGSLVCGTATITATITATFLAGTKFWGALLSFV 256
QY 237 LVVCAITLGDVLESQPKRDIGIKDMNLLPGHGLMDRLDGMPLPAAVWTWILSVI 291
Db 257 LVLTCTLGDVLESQPKRDIGIKDMGRLLPGHGLMDRLDGMPLPAAVWTWILSVI 311
```

```
RESULT 9
US-10-156-761-10160
; Sequence 10160, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: HORIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10160
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10160
```

```
Query Match 35.3%; Score 551; DB 14; Length 392;
Best Local Similarity 39.1%; Pred. No. 3.5e-45;
Matches 117; Conservative 52; Mismatches 114; Indels 16; Gaps 5;

QY 3 EPEQHRSNEMPKP-----KNNAGRDLEKAAIANGVIGLVLLGVLSPNGWYILVAGFM 57
Db 102 QPKPEMPDAPQAPAPQKSAGRDLEKAAIANGVIGLVLLGVLSPNGWYILVAGFM 161
QY 58 AAATWEVGSRLKE--GGYHLPLPIMTIIGQAIWLSPFGTGMILASVFATVILVMYFRIF 116
Db 162 VVGLWELTSRLERKGIKAPLVLAVGGAAMVAVVGRGAGAWAVALTALAVLRM- 220
QY 117 YNTEKEARNYLDTSVGIFVLTWIPFGSFAAMLSMNNNSIPGTFFILTMCLVIASD 176
Db 221 ----TEPPEGYLDVTAGVFAAFVVPFLATFVAMMLTADD-----GPRVLTFLLTWS 272
QY 177 VGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVVTGALSVEHFLDDH--HMMGVILGC 235
Db 273 TGAYATGWRFGXHLAPRISPGKTRGLVGVSPANVAGALCNEFLIDGSGWQGLVGF 332
QY 236 ALVVCATLGDVLESQPKRDIGIKDMNLLPGHGLMDRLDGMPLPAAVWTWILSVISS 294
Db 333 AVAASATLGDGESMIKRDIGIKDMGTLLPGHGGIMDRDLSLLPTAPVVMVLLVIFVGS 391
```

```
RESULT 10
US-10-282-122A-69792
; Sequence 69792, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```

```

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69792
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
; US-10-282-122A-69792

Query Match          19.8%; Score 310; DB 12; Length 271;
Best Local Similarity 33.3%; Pred. No. 7.7e-22;
Matches          99; Conservative          44; Mismatches          84; Indels          70; Gaps          17;

QY 22 RLKAAIANGIGL-GVVLVLG-----IVLSPWGWYILVAGFMAAATWEVG----- 65
DB 5 RIITALLIPALCGFFLLTGTMYFALF-GVVVVLGAWEWARL-AGF-AAQSMRVGYAALV 62
QY 66 SRLKEGGVHLP-LPMTIIGQAIILWSLSPFGTGMGILASPVATVLVLMYFRIFYNGTEKE 123
DB 63 AVLLFFMYLLGLEFWLVA--AVIWS-----VATFLVLY-----PD 99
QY 124 ARNYLEDTS-----VGIFVLTWITPLFGSPAAMLNMONNSIP-GTVFILTFLMCLVADVG 178
DB 100 SSSHASAAACKLVGLLIL-----LPAMQGLVLIKQWPLGNWLLLSVVLVYWAADIG 151
QY 179 GYIAGVFFGSHPWAPLVSPPKSNWEGFAGSIVLG-SVTGALSVMHLLDHHW-----WVGVL 233
DB 152 AYFSGKAFGRKRLAPKVSPPKSNWEGYGLVLSLGITAAVGV-----RDWTVVGFIAALL 207
QY 234 GCALVV-CATLGDLVESQPKDLGKIDMSNLLPGRHGLMDRLDGMPLA-----AMVTW 285
DB 208 GAAVIVFISVIGDLTESMFKRGSKQGVKDSNLLPGRHGLVDRIDSLTAIIPVFAVLW 264

RESULT 11
US-10-282-122A-55991
; Sequence 55991, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55991
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
; US-10-282-122A-55991

Query Match          19.7%; Score 308; DB 12; Length 285;
Best Local Similarity 30.7%; Pred. No. 1.3e-21;
Matches          89; Conservative          47; Mismatches          92; Indels          62; Gaps          13;

QY 30 VGTGLGVLLGLVLSPWG-----WVLLVAGFMAAATWEVGRLEGGVHLP 77
DB 26 VGFAIVTLVVCLAAWEGWQSGFTSRQTVMLVLCGFILAI--MLFTLPE--YHHDI 80
QY 78 PMIIIGQAIILWSLSPFGTGMGILASPVATVLVLMYFRIFYNGTEKEARN--YLRDTSVGI 135
DB 81 HQPLVAGS--LWIS-----LAWVVAALLV---LFYPGSASLWRNSKVLK---LI 122
QY 136 F-VLTWITPLFGSPAAMLN-MQNNSTIPGYFYFILTFLMCLVADVGVIAGVFGSHPWAP 193
DB 123 FGLLTIIPTFFWGMVALRAWHYDENHYSGAIWLLVYVWILVWGDGSGAYMFGKLFGRHKLAP 182
QY 194 LVSPKKSWSGFSIVLGSVTGALSVMHLLDHHWGV-----ILGAL--VUCA 241
DB 183 KVSFGKTMQGFICGLFTAIIIS-----WGVGVWANLEVAAPSILVCSIPAALAS 231
QY 242 TLGDLVESQPKDLGKIDMSNLLPGRHGLMDRLDGMPLAAMVTWILSVI 291
DB 232 VLGDLTESMFKREAGIKDSGLHLPGRHGLVDRIDSLTAAPVPVFAVCLLLV 281

RESULT 12
US-09-864-408A-906
; Sequence 906, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 906
; LENGTH: 79

```

Db 29 GALLIAFVVFAAWEWARLLKGGAGPVIYALVAALVASTELGAGIEQA---REL----- 81  
QY 82 IGGQAIINLSWPFGTWGLASFAVATVLYMYRIFRYNGTEKEARNYLRDTSYGVFLTWI 141  
Db 82 FQAAALFWV-----VAGPFVLLRKPRTLSSQA---WRPFLFLAGIVIFVACW- 124  
QY 142 PLFGSFAAMLSMONNSIPGTFFILTFMLCVIASOVGGYIAGVFGSHPMAPLVSPKXSM 201  
Db 125 -----HALVAARMQ-----GVFPVLSLLLLVWLADIGAYFGKAFGKHKLAPISFGKTW 174  
QY 202 EGPAGSIVLGSVTGALSVMHF-----LLDHHWMMGVILGCA-LVVCATLGDIVESQ 250  
Db 175 EGAIGWLAVMIVAAAATVFLHAFEPFLYSALLAHGAVRTLLATLLVVFVSVDLFESE 234  
QY 251 FKRLGKIDMSNLLPGHGGMDRLDGLMPLAAMVTWIL 288  
Db 235 MKTQAGVKDSGLLPGHGGVLDRLDALLVPLMLLL 272

RESULT 14  
US-10-282-122A-67557  
; Sequence 67557, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 67557  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Pseudomonas putida  
US-10-282-122A-67557

Query Match 19.4%; Score 299; DB 12; Length 271;  
Best Local Similarity 34.0%; Pred. No. 9.2e-21;  
Matches 100; Conservative 39; Mismatches 91; Indels 64; Gaps 15;

TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-864-408A-906  
Query Match 19.6%; Score 306; DB 11; Length 79;  
Best Local Similarity 71.8%; Pred. No. 4.1e-22;  
Matches 56; Conservative 10; Mismatches 12; Indels 0; Gaps 0;  
QY 219 VHFLLDHHWMMGVILGCAVVCATLGDIVESQFKRLDGLKIDMSNLLPGHGGMDRLDGLM 278  
Db 1 VHFLLDHHWMMGVILGCAVVCATLGDIVESQFKRLDGLKIDMSNLLPGHGGMDRLDGLM 60  
QY 279 PAAMVTWILSVISSYP 296  
Db 61 PAASATYMLNAAATMNP 78

RESULT 13  
US-10-282-122A-49018  
; Sequence 49018, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49018  
; LENGTH: 273  
; TYPE: PRT  
; ORGANISM: Burkholderia fungorum  
US-10-282-122A-49018

Query Match 19.4%; Score 303; DB 12; Length 273;  
Best Local Similarity 30.6%; Pred. No. 3.8e-21;  
Matches 85; Conservative 38; Mismatches 97; Indels 58; Gaps 9;  
QY 35 GVILVLCIVLSPWGW-----YILVAGFWAAATWVGSLKXGGVHLPLPIWI 81



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:32:48 ; Search time 16 Seconds  
(without alignments)  
1785.555 Million cell updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNEPEQHRSVRMPKPKNA.....LPAAMVTWLILSVISSYPS 297

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*\*

1: Pir1:\*\*

2: Pir2:\*\*

3: Pir3:\*\*

4: Pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	719	46.0	306	D70924	probable cdaA prot
2	714	45.7	312	G87107	probable phosphati
3	539	34.5	391	TJ35470	probable integral
4	303	19.4	271	JC4832	phosphatidate cyti
5	299	19.1	282	A10128	phosphatidate cyti
6	298	19.1	271	F83188	phosphatidate cyti
7	298	19.1	280	E82099	phosphatidate cyti
8	298	19.1	285	A20529	phosphatidate cyti
9	286.5	18.3	265	A82000	phosphatidate cyti
10	285	18.2	249	1 SYECDG	phosphatidate cyti
11	285	18.2	249	A39651	CDP-diglyceride sy
12	285	18.2	249	A35502	CDP-diglyceride sy
13	280.5	17.9	265	B81229	phosphatidate cyti
14	277	17.7	288	G54102	phosphatidate cyti
15	265.5	17.0	294	A22290	phosphatidate cyti
16	259.5	16.6	284	B75386	phosphatidate cyti
17	253	16.2	399	T47873	phosphatidate cyti
18	252.5	16.2	430	A84887	probable phosphati
19	251.5	16.1	259	A70408	phosphatidate cyti
20	249	15.9	241	H81278	phosphatidate cyti
21	248.5	15.9	278	A83355	phosphatidate cyti
22	246.5	15.8	311	A83329	probable phosphati
23	245.5	15.7	310	H82138	probable phosphati
24	242.5	15.5	277	A82746	phosphatidate cyti
25	242.5	15.5	277	B37527	hypothetical prote
26	240.5	15.4	269	G69597	phosphatidate cyti
27	239.5	15.3	266	G54546	CDP-diglyceride sy
28	238.5	15.3	312	G70114	phosphatidate cyti
29	236.5	15.1	260	H89899	phosphatidate cyti

30	236.5	15.1	298	2	C90880	probable phosphati
31	236.5	15.1	298	2	F85738	probable phosphati
32	235.5	15.1	284	2	F83952	phosphatidate cyti
33	234.5	15.0	284	2	A82729	phosphatidate cyti
34	231.5	14.8	266	2	F71961	cdp-diacylglycerol
35	231	14.8	275	2	A87487	phosphatidate cyti
36	230.5	14.7	262	2	AH1601	phosphatidate cyti
37	230.5	14.7	298	2	D64892	phosphatidate cyti
38	228.5	14.6	262	2	AD1239	phosphatidate cyti
39	225.5	14.4	267	2	A97902	phosphatidate cyti
40	224	14.3	293	2	S77254	phosphatidate cyti
41	223	14.3	230	2	E97773	hypothetical prote
42	219.5	14.0	267	2	G95030	phosphatidate cyti
43	216.5	13.9	228	2	G71700	phosphatidate cyti
44	206	13.2	378	2	E82882	CDP-diglyceride sy
45	204	13.1	287	2	D71304	probable phosphati

ALIGNMENTS

RESULT 1

D70924

Probable cdaA protein - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C;Accession: D70924

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: D70924

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-306 <COL>

A;Cross-references: GB:Z74024; GB:AL123456; NID:G3250700; PIDN:CAA98357.1; PID:gl403401

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: cdaA

C;Superfamily: phosphatidate cytidyltransferase

Query Match	46.0%	Score	719;	DB	2;	Length	306;
Best Local Similarity	47.7%	Pred. No.	2.1e-54;				
Matches	143;	Conservative	46;	Mismatches	97;	Indels	14;
							Gaps 3;
QY	2	NEPEQHRSRM--	PKPKNAGRDLDKAAI	AVGIGLVGLVLS	PGWGYILVAGFMAA	59	
Db	10	NPAEQPARAKQOPAT	ETSRAGDRLRAIV	GLSGLVLI	AVLVFVFRVVAIVAVATLV	69	
QY	60	ATWEVSRLEKEGYHLP	LPIMIGGQAI	IWLSPFGT	MGILASFVATVLVLMYFRIFY--	117	
Db	70	ATHEVVRRLREAGYLIP	IVPLLLIGGQAAV	LWTFPGAV	GALAGFGGMVVCMIWRLFMQD	129	
QY	118	-----NGTEKEARNY	LDTSVGI	FVLTVIP	LEGSPAAMLSLMONNSIDGT	171	
Db	130	SVTRPTTGAPSPGNY	LSVDSATV	FLAVWVPL	FCSEGLVVPEN---GSGWVFCNMIA	185	
QY	172	VIASDVGGYIAGVY	FGSHPMAPLV	SPKSWEGFAGS	IVLGSVTGALSVHFLDHHMMGV	231	
Db	186	VIASDVGGYAVGV	LVFGKHPMVT	ISPKKSWEGFAGS	LVCGITATITATFLVGTKPWIGA	245	
QY	232	ILGCAIVVCATIGD	IVESQFKRDLG	IKOMNSNLL	PHGGMLMDRLDGMPLPAAMVTWLILSVI	291	
Db	246	LLGVFLVLTALGDL	VESQVKRDLGI	KDNGRL	LLPGHGLMDRLDGLTLP	305	

RESULT 2

G87107

probable phosphatidate cytidyltransferase [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
C;Accession: G87107  
R;Cole, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hc  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A;Title: Massive gene decay in the leprosy bacillus.  
A;Reference number: A86909; PMID:21128732; PMID:11234002  
A;Accession: G87107  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-312 <STO>  
A;Cross-references: GB:AL450380; NID:gl3093385; PIDN:CAC30540.1; GSPDB:GN00147  
C;Genetics:  
C;Superfamily: phosphatidate cytidyltransferase  
Query Match 45.7%; Score 714; DB 2; Length 312;  
Best Local Similarity 49.5%; Pred. No. 5.8e-54;  
Matches 146; Conservative 49; Mismatches 84; Indels 16; Gaps 5;  
QY 9 RSMRM-PPKPNAGRDAAIAVIGLGLVLLGIVLSPWGMVILVAGFMMAATWEVGRS 67  
DB 21 RAMRQSTKNTPRAGNLPAATIAVGLSIGVLVATLVFAPRIWVVLCAIAIFVAGHEVVR 80  
QY 68 LKEGGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFY--NGTEKEAR 125  
DB 81 LREAGYVIPAIPLLIGGQFTWLTWTPYTVGALAGFGATVVCWIMIRVMDHNSKQHSR 140  
QY 126 -----NVLRTDSVGIFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMFCVIA 176  
DB 141 EALAGPPVSNVLRDASATVFLAAWPLFASPAALLVPEKGA--GRVFCL--MIAVVASD 196  
QY 177 VGGYTAGVFFGSHPMAPLVSPKKSNEGFAGSVLGSVTGALSVHPLDHHWGMVILGCA 236  
DB 197 VGGYTVGVLFGRHPLVPRISPNKSEGFAGSVLCVCTTATILTATPLAGKTPWVGALLSFV 256  
QY 237 LVVCATLGDVLESQKRDGIKQMSNLLPGHGLMDRLDGMPLPAAMVTLILSVI 291  
DB 257 LVLTCTGLDVSQKRDGIKQMSNLLPGHGLMDRLDGMPLPAAMVTLILSVI 311  
RESULT 3  
T35470  
probable integral membrane protein - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1998  
A;Reference number: Z21578  
A;Accession: T35470  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-391 <MUR>  
A;Cross-references: EMBL:AL031035; PIDN:CAA19924.1; GSPDB:GN00070; SCODEB:SC6A9.39c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCODEB:SC6A9.39c  
Query Match 34.5%; Score 539; DB 2; Length 391;  
Best Local Similarity 38.8%; Pred. No. 8.2e-39;  
Matches 114; Conservative 51; Mismatches 117; Indels 12; Gaps 5;  
QY 4 PQQHRSMPKP-KNNAGRDAAIAVIGLGLVLLGIVLSPWGMVILVAGFMMAATW 62  
DB 107 PDAQPQSQPPQKKSAGRDGAIGVGLGVIIASLFVVKAVFGVTAIVAVVGLW 166  
QY 63 EYGSRLKE-GGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFYNGTE 121  
DB 167 ELTKSLERKGIKAPLVPLAIGGAAMVAGYARGAEGAWAMALTALAVLVRM-----T 221

QY 122 KEARNYLRTSVGIFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMFCVIA 181  
DB 222 EPPEGYLDKVTAGLFAAFVPLAFVAMLAADD-----GAWRVLFLLTLTVSTGTAYA 277  
QY 182 AGVFFGSHPMAPLVSPKKSNEGFAGSVLGSVTGALSVHFLDLH-HWMMGVILGALVVC 240  
DB 278 VGMRFGRKKLAPRISPGKTREGLLGAIAFAMVAGALCMQFLIDGAWMQGLLGLWAVS 337  
QY 241 ATLGDLVSESQKRDGIKQMSNLLPGHGLMDRLDGMPLPAAMVTLILSVISSS 294  
DB 338 ATGLDGLGESMIKRDGIKQMSNLLPGHGLMDRLDGLLPTAPVWVLLLVIFVGS 391  
RESULT 4  
JC4832  
phosphatidate cytidyltransferase (EC 2.7.7.41) - Pseudomonas aeruginosa  
N;Alternate names: CDP-diglyceride synthetase  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 20-Jun-2000  
C;Accession: JC4832  
R;Taguchi, K.; Fukutomi, H.; Kuroda, A.; Kato, J.; Ohtake, H.  
Gene 172, 165-166, 1996  
A;Title: Cloning of the Pseudomonas aeruginosa gene encoding CDP-diglyceride synthetase.  
A;Reference number: JC4832; PMID:96257274; PMID:8654980  
A;Accession: JC4832  
A;Molecule type: DNA  
A;Residues: 1-271 <TAG>  
A;Cross-references: DDJ7:D50811; NID:gl262331; PIDN:BA049437.1; PID:gl262332  
C;Comment: This enzyme catalyses the condensation of CTP and phosphatidic acid to form C1  
C;Genetics:  
C;Superfamily: phosphatidate cytidyltransferase  
C;Keywords: nucleotidyltransferase  
Query Match 19.4%; Score 303; DB 2; Length 271;  
Best Local Similarity 33.3%; Pred. No. 1.1e-18;  
Matches 99; Conservative 39; Mismatches 93; Indels 64; Gaps 15;  
QY 22 RDLKAAIAVIGLGLVLL-----GIVLS--PWGYILVAGF-----MAAATWE 63  
DB 5 RIITALLVPLALGGFFLLEGAFALFICAVVSLGAWENARL-AGVEQOQGRVAVATAVA 63  
QY 64 VGSRLKEGGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFYNGTEKE 123  
DB 64 V---LMVALYHLP---QLAGAVLLGLAVW-----WTLATVLTLY-----PE 99  
QY 124 ARNYL--RDTSVGIFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMFCVIA 181  
DB 100 SVGYWGRWRRLGMLLILLPAWGLVLLKQVAAANGL-----IIAVVVLWGADIGAYF 154  
QY 182 AGVFFGSHPMAPLVSPKKSNEGFAGSVLGSVTGALSVHFLDHHWMMG-----VILGCA 236  
DB 155 SGKAFGRKKLAPRISPGKTREGLLGAIAFAMVAGALCMQFLIDGAWMQGLLGLWAVS 210  
QY 237 LVVCATL-GDLVESQKRDGIKQMSNLLPGHGLMDRLDGMPLA-----AMVTW 285  
DB 211 LVVTVSIVGDLIESMPKRSKIKSSNLLPGHGLVLDRLDLSLTAAIPVFAALLW 264  
RESULT 5  
AI0128  
phosphatidate cytidyltransferase (EC 2.7.7.41) [imported] - Yersinia pestis (strain CO  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AI0128  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; PMID:21470413; PMID:11586360  
A;Accession: AI0128  
A;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-282 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC98982.1; PID:g15979117; GSPDB:GN00175

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F83188

A:Status: preliminary

A:Residues: 1-271 <STO>

A:Cross-references: GB:AB004785; GB:AE004091; NID:G9949809; PIDN:AAG07039.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: cdaA; PA3651

C:Superfamily: phosphatidate cytidylyltransferase

Query Match 19.1%; Score 299; DB 2; Length 282;

Best Local Similarity 29.6%; Pred. No. 2.5e-18;

Matches 84; Conservative 47; Mismatches 85; Indels 68; Gaps 12;

QY 32 IGLGVLLGLVLSWGW-----YILVAGFMAAATWEVGSRLKEG 71

DB 26 VGFALVILVCLAAWEGQLAGPASRTQRIWAILCGFLVAMLSLPEYQ----- 77

QY 72 GYHPLPIMIGGQAIWLS---WPGFMGILASGFATVVLVLMYFRIFYNGTEKEARNYL 128

DB 78 --HSPHLLV--STPLSLSGMW-----VAALMLVLT-----PRSAVSWR 114

QY 129 RDTSGVIF--VLTWIPLF--GSFAAMLSLMQNNISPGTYFILTFLMTCVIASDVGGYIAGVF 185

DB 115 NSRLRIIFGLTIIPFPWGFALRQYGYEQNHNTGAWLLYVLLVWGADSGAYMFGKL 174

QY 186 FGSHPMAPLVSPKKSWEFGFAGSIVLGSVTGALSVHFLLDHHWGMVI---LGCALV--V 239

DB 175 FGKHLAPKVSFGKTWEGLIG---GLITSAL--ISLFGRYAPLDIVPEKLLICSVVAAL 229

QY 240 CATIGDLVESQKRDGLGKDNMNLPGHGLMDRLDGLMPAMV 283

DB 230 ASVLGDLTESMFKREAGIKDSGLHPGSGILDRIDSLTAAPV 273

RESULT 6

F83188

Phosphatidate cytidylyltransferase PA3651 [imported] - Pseudomonas aeruginosa (strain PA

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F83188

A:Status: preliminary

A:Residues: 1-271 <STO>

A:Cross-references: GB:AB004785; GB:AE004091; NID:G9949809; PIDN:AAG07039.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: cdaA; PA3651

C:Superfamily: phosphatidate cytidylyltransferase

Query Match 19.1%; Score 298; DB 2; Length 271;

Best Local Similarity 33.3%; Pred. No. 2.9e-18;

Matches 98; Conservative 38; Mismatches 94; Indels 64; Gaps 15;

QY 22 RDLKAAIYAVGIGLVLL-----GIVLS--PWGWIYLVAGF-----MAAATWE 63

DB 5 RIITAVLLPALTGGFFLEGAFFALTGAVVSLGANEWAPL--AGYEQGRVAYATAVA 63

QY 64 VGSRLKEGGYHLPIMIIIGQAIWLSWPGFMGILASGFATVVLVLMYFRIFYNGTEKE 123

DB 64 V---LWVALYHLP---QLAGAVLLALVW-----WTLATVLT-----PE 99

QY 124 ARNYL--RDTSGVIFVLTWIPLF--GSFAAMLSLMQNNISPGTYFILTFLMTCVIASDVGGYI 181

DB 100 SVGWGGRWRRLGMLLILPAWQGLVLLKQWPLANGL-----IYAVVLVWGDIGAYF 154

QY 182 AGVFFGSHPMAPLVSPKKSWEFGFAGSIVLG--SVTCALSVHFLLDHHWGMV-----VILGCA 236

DB 182 AGVFFGSHPMAPLVSPKKSWEFGFAGSIVLG--SVTCALSVHFLLDHHWGMV-----VILGCA 236

Db 155 SGKAFKRLAPRVSPKSGWGVYGLAASLAITLAVGLY----RGWSLGLALLLGLAA 210

QY 237 LVVCATL--GDLVESQKRDGLGKDNMNLPGHGLMDRLDGLMPLA-----AMVTM 285

Db 211 LVVFSIVGDLTSMFKRQSGIKSSNLLPGHGVLDRLDLSLTAAPVFAALLW 264

RESULT 7

E82099

Phosphatidate cytidylyltransferase VC2255 [imported] - Vibrio cholerae (strain N16961 ser

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: E82099

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bais, S.; Qin, H.; Dragoi, I.; Sellers, P.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82099

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <HEI>

A:Cross-references: GB:AE004297; GB:AE003852; NID:G9656810; PIDN:AAF95399.1; GSPDB:GN001;

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2255

A:Map position: 1

C:Superfamily: phosphatidate cytidylyltransferase

Query Match 19.1%; Score 298; DB 2; Length 280;

Best Local Similarity 29.7%; Pred. No. 3e-18;

Matches 90; Conservative 41; Mismatches 106; Indels 66; Gaps 10;

QY 24 LKAAIYAVGIGLVLLGLVLSWGWYILVAGFMAAAT-----WEVGSRLKEGGYHLP-LP 78

Db 1 MKQRIITAILAPLVILGILYLPFAWFLA---LAVVTLGFWMTQFVNQPSRLMAMIP 57

QY 79 IMIIGQAIWLSWPGFMGILASGFATVVLVLMYFRIFYNGTEKEAR 125

Db 58 ALLVAGISVALIDFQFPAISNNNTAFAIVLGSLWLVSSGLAITPR-----SR 108

QY 126 NYLRDTSV-----GIFVLT---WIPFGSFAAMLSLMQNNISPGTYFILTFLMTCVIASDV 177

Db 109 PLWEHSSTVRHLFGTLTLLPFFVSVLFLRADTVLS----DPLYGAKLVLFVCLVWAADS 164

QY 178 GGYIAGVFFGSHPMAPLVSPKKSWEFGFAGSIVLGSVTGALSVHFLLDHHWGMVILGCA 237

Db 165 GAYFVGSIGKHKMAPAVSPNKTIBGLVGGIVTAMLVG-----YVVAECFGIQF 213

QY 238 -----VVCATIGDLVESQKRDGLGKDNMNLPGHGLMDRLDGLMPLAAMVT 284

Db 214 SSMPVMLLIILLTVVLSVLGDLVESMFKRQSGIKSSNIIIPGHGILDRIDSLTAAPV 273

QY 285 WLI 287

Db 274 ALL 276

RESULT 8

AH0529

Phosphatidate cytidylyltransferase [imported] - Salmonella enterica subsp. enterica serov

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AH0529

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, E.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608





```

RESULT 11
A99651
CDP-diglyceride synthetase [imported] - Escherichia coli (strain O157:H7, substrain RIMD
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C/Accession: A99651
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: A99651
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-249 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA833600.1; PID:g13359633; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs0177
C/Superfamily: phosphatidate cytidyltransferase

Query Match 18.2%; Score 285; DB 2; Length 249;
Best Local Similarity 31.0%; Pred. No. 3.4e-17;
Matches 85; Conservative 38; Mismatches 87; Indels 64; Gaps 12;

QY 47 WG-----WYILVAGFMAAATWEVGSRLKEGGYHLPIMIIIGQAIIWLS-- 91
DB 7 WQQLSGFTTRSORVWLAIVLCGLLLALMLFL---LPEYHRNIHQPLVEIS 59
QY 92 WPFGTMGILASFVATVLMVYFRIFYNGTEKEARNYLRDTSVGIF-VLTWIPLF-GSPAA 149
DB 60 W-----IVALLLVL-----FYPGSAATWRN--SKTLRLIFGVLTIVPFPGMLAL 102
QY 150 MSLMQNNSIPGTYFTILTMCLVIAVDVCGYTAGVFFGSHPMAPLVSPKKSWEFGAGSIV 209
DB 103 RAWHYDENHYSGAIIWLYVWLVWGADSGAYMFGKLFQGHKLAPKVPSPKWTGQFIGNLA 162
QY 210 LGSVTGALS VHFLDHHWMMGV-----ILGCALV--VCATIGDLVSESQFKRDLGI 257
DB 163 TAAVIS-----WGYGMANLDAVPVTLICSIVAALASVLGDLTESMFKREAGI 211
QY 258 KMSNLLPGHGLMDRLDGMPLPAAMVTWLLSVI 291
DB 212 KDSGHLIPGHGILDRIDSLTAAPVPFACLLLV 245

RESULT 12
A85502
CDP-diglyceride synthetase [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: A85502
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: A85502
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-249 <STO>
A/Cross-references: GB:AB005174; NID:g12512900; PIDN:AAG54477.1; GSPDB:GN00145; UWGP:201
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: cdeA
C/Superfamily: phosphatidate cytidyltransferase

Query Match 18.2%; Score 285; DB 2; Length 249;
Best Local Similarity 31.0%; Pred. No. 3.4e-17;
Matches 85; Conservative 38; Mismatches 87; Indels 64; Gaps 12;

QY 47 WG-----WYILVAGFMAAATWEVGSRLKEGGYHLPIMIIIGQAIIWLS-- 91
DB 7 WQQLSGFTTRSORVWLAIVLCGLLLALMLFL---LPEYHRNIHQPLVEIS 59
QY 92 WPFGTMGILASFVATVLMVYFRIFYNGTEKEARNYLRDTSVGIF-VLTWIPLF-GSPAA 149
DB 60 W-----IVALLLVL-----FYPGSAATWRN--SKTLRLIFGVLTIVPFPGMLAL 102
QY 150 MSLMQNNSIPGTYFTILTMCLVIAVDVCGYTAGVFFGSHPMAPLVSPKKSWEFGAGSIV 209
DB 103 RAWHYDENHYSGAIIWLYVWLVWGADSGAYMFGKLFQGHKLAPKVPSPKWTGQFIGNLA 162
QY 210 LGSVTGALS VHFLDHHWMMGV-----ILGCALV--VCATIGDLVSESQFKRDLGI 257
DB 163 TAAVIS-----WGYGMANLDAVPVTLICSIVAALASVLGDLTESMFKREAGI 211
QY 258 KMSNLLPGHGLMDRLDGMPLPAAMVTWLLSVI 291
DB 212 KDSGHLIPGHGILDRIDSLTAAPVPFACLLLV 245

```

```

DB 7 WQQLSGFTTRSORVWLAIVLCGLLLALMLFL---LPEYHRNIHQPLVEIS-----LWASLWG 59
QY 92 WPFGTMGILASFVATVLMVYFRIFYNGTEKEARNYLRDTSVGIF-VLTWIPLF-GSPAA 149
DB 60 W-----IVALLLVL-----FYPGSAATWRN--SKTLRLIFGVLTIVPFPGMLAL 102
QY 150 MSLMQNNSIPGTYFTILTMCLVIAVDVCGYTAGVFFGSHPMAPLVSPKKSWEFGAGSIV 209
DB 103 RAWHYDENHYSGAIIWLYVWLVWGADSGAYMFGKLFQGHKLAPKVPSPKWTGQFIGNLA 162
QY 210 LGSVTGALS VHFLDHHWMMGV-----ILGCALV--VCATIGDLVSESQFKRDLGI 257
DB 163 TAAVIS-----WGYGMANLDAVPVTLICSIVAALASVLGDLTESMFKREAGI 211
QY 258 KMSNLLPGHGLMDRLDGMPLPAAMVTWLLSVI 291
DB 212 KDSGHLIPGHGILDRIDSLTAAPVPFACLLLV 245

RESULT 13
B81229
Phosphatidate cytidyltransferase NMB0185 [imported] - Neisseria meningitidis (strain MC
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: B81229
R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: B81229
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-265 <PET>
A/Cross-references: GB:AE002375; GB:AB002098; NID:g7225394; PIDN:AAF0642.1; PID:g7225404
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB0185
C/Superfamily: phosphatidate cytidyltransferase

Query Match 17.9%; Score 280.5; DB 2; Length 265;
Best Local Similarity 28.9%; Pred. No. 9e-17;
Matches 83; Conservative 50; Mismatches 113; Indels 41; Gaps 9;

QY 24 LKAAIAGVIGLGVLLGIVLSP--KGWYILVAGFMAAATWE---VGSRLKEGGYHLPPI 79
DB 2 LKQRTVITAWLLPLMLGLMDFYAPQWLMAAFQGLIALIALWEYARMGGLCKIKTNHYLAAT 61
QY 80 MIIGQA-----IIWLSWPFMTGILASFVATVLMVYFRIFYNGTEKEARNYLRD 130
DB 62 LVFGVVAYAGGWMLPNLYVYV-----VLAPWLAVMPLWLRFRKWLNGGWQ----- 106
QY 131 TSVGVFLVTWIPLFSGFAAMLSIM--QNNISIPGTYFTILTMCLVIAVDVCGYTAGVFFGS 188
DB 107 ----VYAVGMLLWMPFWFALVSLRPHPDALP-----LLAVMGLVWVADICAFYSGKAFGK 158
QY 189 HPWAPLVSPKKSWEFGAGSIVLGSV--TGALSVHFL-LDHHWMMGVILGCALVWCATIGD 245
DB 159 HKIAPAFSGKSWEGAAGVAVYVMTAVRAGWLAFTGTFDVLIGLVTIVVSCGD 218
QY 246 LVESQFKRDLGKDNKNLLPGHGLMDRLDGMPLPAAMVTWLLSVIS 292
DB 219 LLESMLKRAAGIKDSSKLLPGHGGVFDRTDLSLIAVIVSYAAMMSVLN 265

RESULT 14
G64102
Phosphatidate cytidyltransferase homolog - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 21-Jan-2000
C/Accession: G64102

```

61 QY -----TWEVGSRLKEGGYHLPIMIGGQAIWLSPWPTGMLASFTVATVLVL 110  
 58 Db TLFLSQVLLVTCITLDSNLVDA-----VMPIAGTFICFYLLFQPKMATIADVSASIMGL 110  
 111 QY MYFRIFYNGTEKEARNYLDTSVGIFVLTWIPLFGSFAVLS--LMQNN--SIPCTY--F 164  
 111 Db FVGYL-----PSYVWLRSIGSTSISNPLGGYWPGSITDILEERNFASLFCQFKIT 163  
 165 QY ILTFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKKSWEQ--FAGSIVLG-SVTGALS-VH- 220  
 164 Db VLTFF-LCIWAADIGAYTIGKFFGKTRLS-ISPKTVEGAVEGISASLAVAGAFYLHL 221  
 221 QY ---FLDHHWMMGVILGCALVVCATLGDIVESQFKDLGIKMSNLLPGHGLMDRLDQML 278  
 222 Db POFLL-----TGTILGLLIGLASLLGLTESMLKRDAGVQDSGQLIPGHGGLDRTDSYI 276  
 279 QY PAAMVTWLILSVI 291  
 277 Db FTAPLVYYFVTLI 289

Search completed: August 11, 2004, 13:36:26  
 Job time : 16 secs

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
 Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630; PMID:7542800  
 A:Accession: G64102  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-288 <NIGR>  
 A:Cross-references: GB:U32773; GB:L42023; NID:G1573932; PIDN:AAC22577.1; PID:G1573940;  
 C:Superfamily: phosphatidate cytidylyltransferase

Query Match 17.7%; Score 277; DB 2; Length 288;  
 Best Local Similarity 28.9%; Pred. No. 2e-16;  
 Matches 87; Conservative 32; Mismatches 76; Indels 106; Gaps 13;

QY 32 IGLGVLLVILGIVLSPWGH-----YILVAGFMAA--ATWEVGSRLKEGGY----- 73  
 Db 30 LALGAVAILGI-----WENTQFARLKQPLRPFVTFGLGVFFLW-----LYTEGNYLDAGR 81  
 QY 74 ---HLPPLMIGQAIW-----LSPW-----FQTMGILASFV 104  
 Db 82 VPEQHLQLLI-----NAVSWGLALLLVISYPSKAKFWSKNPLQLLFAFST---LIPFV 134  
 QY 105 ATVLVIMYFRIFYNGTEKEARNYLDTSVGIFVLTWIPLFGSFAVLSLMONNSIPCTYF 164  
 Db 135 AGVLR-----RLEHYTHDPYHGLFLLLV-----F 160  
 QY 165 ILTFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKKSWEQFAGSIVLGSVTGALS-VHF--- 221  
 Db 161 IL-----VWAADSGAYFGRAPGKRLAPKVPKSGKWEVGGTLITVLVLAIFIHFSNN 215  
 QY 222 ---LDPHMMWMMGVILGCALVVCATLGDIVESQFKDLGIKMSNLLPGHGLMDRLDQMLP 279  
 Db 216 TLVGRNITGFIILSVATVATVLSVLGDLTFESMFKRESGVKDSQLIPGHGGLDRLDQMLP 275  
 QY 280 A 280  
 Db 276 A 276

RESULT 15  
 AD2290  
 phosphatidate cytidylyltransferase [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AD2290  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
 Nakazaki, N.; Shimoto, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AD2290  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-294 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA075574.1; PID:G17133009; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all3875  
 C:Superfamily: phosphatidate cytidylyltransferase

Query Match 17.0%; Score 265.5; DB 2; Length 294;  
 Best Local Similarity 29.4%; Pred. No. 2e-15;  
 Matches 92; Conservative 49; Mismatches 101; Indels 71; Gaps 15;

QY 22 RDLKAAIAGVIGLGVLLVILGIVLSPWGHYILV-----AGFMAA-- 60  
 Db 5 RIISGIVAILAL-VAVLLG-----GWYFTIMLAIIVVLGQBYFDIVRTRGILPAKT 57

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:28:52 ; Search time 14 Seconds  
(without alignments)  
1104.631 Million cell updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNEFEQHRSRMKPKNNA.....LPAAWTVLILSVSSYPS 297

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : .SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	719	46.0	306	1 CDSA_MYCTU	Q10807 m phosphati
2	714	45.7	312	1 CDSA_MYCLE	Q9cbul m phosphati
3	298	19.1	271	1 CDSA_PSEAE	Q59640 p phosphati
4	285	18.2	249	1 CDSA_ECOLI	P06466 p phosphati
5	277	17.7	288	1 CDSA_HAEIN	P44937 h phosphati
6	251	16.1	259	1 CDSA_AQUAE	O87292 a phosphati
7	244	15.6	270	1 CDSA_BRUAB	Q89173 b phosphati
8	243.5	15.4	269	1 CDSA_BACSU	Q31752 b phosphati
9	240.5	15.3	266	1 CDSA_HELPY	Q25004 h phosphati
10	239.5	15.3	266	1 CDSA_HELPY	Q25004 h phosphati
11	231.5	14.8	266	1 CDSA_HELPY	Q25004 h phosphati
12	230.5	14.7	298	1 YNBB_ECOLI	P76091 escherichia
13	224	14.3	293	1 CDSA_SYNY3	P73548 s phosphati
14	216.5	13.9	228	1 CDSA_RICHP	Q92d48 r phosphati
15	202	12.9	270	1 CDSA_THENA	Q9xib7 t phosphati
16	201.5	12.9	447	1 CDSA_DROME	P56079 t phosphati
17	198.5	12.7	445	1 CDSA_HUMAN	Q95674 h phosphati
18	197.5	12.6	424	1 CDS1_SOLTU	O04940 s phosphati
19	197	12.6	421	1 CDS1_ARATH	O04928 a phosphati
20	195.5	12.5	461	1 CDS1_HUMAN	Q22903 a phosphati
21	194.5	12.4	461	1 CDS1_RAT	Q35052 r phosphati
22	193.5	12.4	308	1 CDSA_CHLPN	Q9z7y6 c phosphati
23	188.5	12.1	305	1 CDSA_CHLTR	O84457 c phosphati
24	180.5	11.5	305	1 CDSA_CHLMU	Q9pjul c phosphati
25	177	11.3	455	1 CDS1_CAEEL	P53439 caenorhabdi
26	169.5	10.8	395	1 CDSA_MYCPN	P75160 m putative
27	161	10.3	305	1 CDSA_MYGE	Q49433 m putative
28	142	9.1	457	1 CDS1_YEAST	P38221 s phosphati
29	119	7.6	136	1 CDS1_MOUSE	P98191 m phosphati
30	115.5	7.4	326	1 MRAY_STRFP	Q9zha5 streptococc
31	115.5	7.4	326	1 MRAY_STRFP	Q9zha5 streptococc
32	113	7.2	417	1 CLCB_SALTI	Q8z6v0 salmonella
33	112.5	7.2	329	1 MRAY_LACLA	Q9ch70 lactococcus

#### ALIGNMENTS

##### RESULT 1

ID	CDSA_MYCTU	STANDARD	PRT	306 AA
AC	Q10807			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG synthase) (CDP-DG synthetase)			
GN	CDSA OR RV2881C OR MT2948 OR MTCY274.12C OR MB2905C.			
OS	Mycobacterium tuberculosis, and			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773, 1765;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.tuberculosis; STRAIN=H37Rv;			
RX	MEDLINE=98295987; PubMed=9634230;			
RA	Cole S.T., Broech R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream V.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;			
RA	"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";			
RL	Nature 393:537-544 (1998).			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;			
RX	MEDLINE=22206494; PubMed=12218036;			
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;			
RA	"Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";			
RL	J. Bacteriol. 184:5479-5490 (2002).			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.bovis; STRAIN=AF2122/97;			
RX	MEDLINE=22709107; PubMed=12788972;			
RA	Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H., Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S., Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R., Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;			
RA	"The complete genome sequence of Mycobacterium bovis.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).			
CC	-!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-diacylglycerol.			

Q58004 methanococc  
Q55986 synchocyst  
Q8Y5M0 listeria m  
Q8ZBK5 salmonella  
Q929Y0 listeria in  
P56834 zymomonas m  
Q9KJ23 pseudomonas  
Q28481 macaca fasc  
P32703 escherichia  
Q97CB5 thermoplasma  
Q13336 homo sapien  
Q43246 homo sapien

34 112 7.2 365 1 Y584 METJA  
35 109 7.0 365 1 MRAY\_SYNY3  
36 107.5 6.9 324 1 MRAY\_LISMO  
37 107 6.8 417 1 CLCB\_SALTY  
38 104.5 6.7 324 1 MRAY\_LISIN  
39 102.5 6.6 356 1 MRAY\_ZYMO  
40 102.5 6.6 578 1 DSHD\_PSESP  
41 101.5 6.5 416 1 RHL\_VACFA  
42 100.5 6.4 549 1 YJCE\_ECOLI  
43 100 6.4 183 1 Y187\_THEVO  
44 100 6.4 389 1 UT1\_HUMAN  
45 99 6.3 636 1 CTR4\_HUMAN

CC -!- PATHWAY: Phospholipid biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the CDS family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z74024; CA98357.1; -;  
 CC EMBL; AB007118; AAK47273.1; ALT INIT.  
 CC EMBL; BX248344; CAD96592.1; -;  
 CC PIR; D70924; D70924.  
 CC TIGR; W2948; -;  
 CC Tuberculin; Rv2881c; -;  
 CC InterPro; IPR000374; PC\_trans.  
 CC Pfam; PF01148; CTP\_transf\_1; 1.  
 CC PROSITE; PS01315; CDS; 1.  
 CC Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;  
 CC Transmembrane; Complete proteome.  
 CC TRANSMEM 36 56 POTENTIAL.  
 CC TRANSMEM 82 102 POTENTIAL.  
 CC TRANSMEM 103 123 POTENTIAL.  
 CC TRANSMEM 151 171 POTENTIAL.  
 CC TRANSMEM 180 200 POTENTIAL.  
 CC TRANSMEM 218 238 POTENTIAL.  
 CC TRANSMEM 241 261 POTENTIAL.  
 CC TRANSMEM 285 305 POTENTIAL.  
 CC SEQUENCE 306 AA; 32035 MW; 8FA81A108035099F CRC64;

DR EMBL; Z74024; CA98357.1; -;  
 DR EMBL; AB007118; AAK47273.1; ALT INIT.  
 DR EMBL; BX248344; CAD96592.1; -;  
 DR PIR; D70924; D70924.  
 DR TIGR; W2948; -;  
 DR Tuberculin; Rv2881c; -;  
 DR InterPro; IPR000374; PC\_trans.  
 DR Pfam; PF01148; CTP\_transf\_1; 1.  
 DR PROSITE; PS01315; CDS; 1.  
 DR Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;  
 DR Transmembrane; Complete proteome.  
 DR TRANSMEM 36 56 POTENTIAL.  
 DR TRANSMEM 82 102 POTENTIAL.  
 DR TRANSMEM 103 123 POTENTIAL.  
 DR TRANSMEM 151 171 POTENTIAL.  
 DR TRANSMEM 180 200 POTENTIAL.  
 DR TRANSMEM 218 238 POTENTIAL.  
 DR TRANSMEM 241 261 POTENTIAL.  
 DR TRANSMEM 285 305 POTENTIAL.  
 DR SEQUENCE 306 AA; 32035 MW; 8FA81A108035099F CRC64;

Query Match 46.0%; Score 719; DB 1; Length 306;  
 Best Local Similarity 47.7%; Pred. NO. 2.5e-47;  
 Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY 2 NEPEQHRSMEM--PKPKNNAGRDLLKAAIAVIGLGLVLLGVLSPGWVILVAGFMAA 59  
 DB 10 NPAEQPARAKQOPATETSRAGRDLLKAAIAVIGLGLVLLGVLSPGWVILVAGFMAA 69  
 QY 60 ATWEVGSRLKEGGYHLPIMIGGQAIILWSWPGTNGILASVATVILVLMYPRIFY-- 117  
 DB 70 ATHEVVRRLREAGYLPVPLIGGQAAVLLTWPGCAVAGALGFGGVMVCMVRLFMQD 129  
 QY 118 -----NGTEKARYLRDTSVGIIVLWIPFGSFAAMLMLQNNISIPGYFILTPLMC 171  
 DB 130 SVTRPTTGGAPSPGNLYSDVSATVFLAVVPLFCFGALVYPEN-----GSGWFCMMIA 185  
 QY 172 VIASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSIVHFLDHHMMGV 231  
 DB 186 VIASDVGGYAVGVLGKHPMVLTISPKKSWEFGAGSLVCGTATITATFLVGKTPMIGA 245  
 QY 232 ILGALVVCATIGDLVESQFKRDIGIKDMSNLLPHGGIMDRDLGMLPAAMVTLISVI 291  
 DB 246 LLGVLFVLTALGDLVESQVKRDIGIKDMSNLLPHGGIMDRDLGMLPAAMVTLISVI 305

RESULT 2  
 CDS\_MYCLE STANDARD; PRT; 312 AA.  
 AC Q9CBUL;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
 DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
 DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG  
 DE synthase) (CDS OR MLI589).  
 GN CDS OR MLI589.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Bigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 Wreeeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 Holroyd S., Hornsby T., Jagels K., Lagroix C., Maclean J., Moule S.,  
 Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 Barrell B.G.;  
 RA "Massive gene decay in the leprosy bacillus.";  
 RT Nature 409:1007-1011(2001).  
 RL Nature 409:1007-1011(2001).  
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
 CC diacylglycerol.  
 CC -!- PATHWAY: Phospholipid biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the CDS family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL583922; CAC30540.1; -;  
 CC PIR; G87107; G87107.  
 CC Leproma; MLI589; -;  
 CC InterPro; IPR000374; PC\_trans.  
 CC Pfam; PF01148; CTP\_transf\_1; 1.  
 CC PROSITE; PS01315; CDS; 1.  
 CC Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;  
 CC Transmembrane; Complete proteome.  
 CC TRANSMEM 37 57 POTENTIAL.  
 CC TRANSMEM 58 78 POTENTIAL.  
 CC TRANSMEM 85 105 POTENTIAL.  
 CC TRANSMEM 110 130 POTENTIAL.  
 CC TRANSMEM 157 177 POTENTIAL.  
 CC TRANSMEM 186 206 POTENTIAL.  
 CC TRANSMEM 223 243 POTENTIAL.  
 CC TRANSMEM 247 267 POTENTIAL.  
 CC SEQUENCE 312 AA; 32972 MW; 5A1C1DFEDD3E4913 CRC64;

Query Match 45.7%; Score 714; DB 1; Length 312;  
 Best Local Similarity 49.5%; Pred. NO. 6.1e-47;  
 Matches 146; Conservative 49; Mismatches 84; Indels 16; Gaps 5;

QY 9 RSMRM--PKPKNNAGRDLLKAAIAVIGLGLVLLGVLSPGWVILVAGFMAAATWEVGS 67  
 DB 21 RANQSTKTPRAGRLPRAIAVIGLGLVLLGVLSPGWVILVAGFMAAATWEVGS 80  
 QY 68 LKEGGYHLPIMIGGQAIILWSWPGTNGILASVATVILVLMYPRIFY--NGTEKAR 125  
 DB 81 LREAGYVIPAIPLLIGGQFTVLTWPTVGTALGAGFATVAVVCMVRLVMDNSQKHSR 140  
 QY 126 -----NYLRDTSVGIIVLWIPFGSFAAMLMLQNNISIPGYFILTPLMCVIA 176  
 DB 141 EALAGPPVSNLYRDSATVFLAAWVPLFASFAALLVYPKDGA--GEVFL--MIAV 196  
 QY 177 VGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSIVHFLDHHMMGVILG 236  
 DB 197 VGGYTVGVLPFGKHPVLPISPNKKSWEFGAGSLVCGTATITATFLAGKTPMVGALLSF 256  
 QY 237 LVVCATIGDLVESQFKRDIGIKDMSNLLPHGGIMDRDLGMLPAAMVTLISVI 291  
 DB 257 LVLTCTIGDLVESQVKRDIGIKDMSNLLPHGGIMDRDLGMLPAAMVTLISVI 311

Query Match 45.7%; Score 714; DB 1; Length 312;  
 Best Local Similarity 49.5%; Pred. NO. 6.1e-47;  
 Matches 146; Conservative 49; Mismatches 84; Indels 16; Gaps 5;

RESULT 3  
CDS: PSEAE  
ID CDS: PSEAE STANDARD; PRT: 271 AA.  
AC Q59640;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
synthase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
synthase) (CDS) (CTP-phosphatidate cytidyltransferase) (CDP-DAG  
synthase) (CDP-DG synthetase).  
GN CDS OR CDS OR PA3651.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
CC Pseudomonadaceae; Pseudomonas.  
CX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=96257274; PubMed=8654980;  
RA Taguchi K., Fukutomi H., Kuroda A., Kato J., Ohtake H.;  
RT "Cloning of the Pseudomonas aeruginosa gene encoding CDP-diglyceride  
synthetase.";  
RL Gene 172:165-166(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.;  
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Gaber R.L., Grolty L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
diacylglycerol.  
CC -!- PATHWAY: Phospholipid biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -!- SIMILARITY: Belongs to the CDS family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D50811; BRA09437.1; -.  
DR EMBL; AB004785; NAG07039.1; -.  
DR PIR; F83188; F83188.  
DR PIR; JC4832; JC4832.  
DR InterPro; IPR000374; PC trans.  
DR Pfam; PF01148; CTP trans1.1.  
DR PROSITE; PS01315; CDS; 1.  
KW Transferase; Nucleoside transferase; Phospholipid biosynthesis;  
KW Transmembrane; Inner membrane. Complete proteome.  
FT TRANSMEM 12 32 POTENTIAL.  
FT TRANSMEM 53 73 POTENTIAL.  
FT TRANSMEM 75 95 POTENTIAL.  
FT TRANSMEM 111 131 POTENTIAL.  
FT TRANSMEM 136 156 POTENTIAL.  
FT TRANSMEM 174 194 POTENTIAL.  
FT TRANSMEM 199 219 POTENTIAL.  
FT TRANSMEM 251 271 POTENTIAL.  
FT TRANSMEM 131 133 WPL -> VAA (IN REF. 1).  
SQ SEQUENCE 271 AA; 28856 MW; 5025059C3F1A64C7 CRC64;

Query Match 19.1%; Score 298; DB 1; Length 271;  
Best Local Similarity 33.3%; Pred. No. 1.2e-15;

Matches 98; Conservative 38; Mismatches 94; Indels 64; Gaps 15;  
QY 22 RDLKRAIAVIGIGLVLL-----CIVLS--PWGMYILVAGF-----MAAATWE 63  
DB 5 RIITALLPILALGGFFLEGAFFALFICAVVSLGAEWARL-AGVEQFGRAVAATVA 63  
QY 64 VSRLEKGGYHPLPMTIIGQAATLWSWFFGTGMLILASFAVATVLMYFRFYNGTEKE 123  
DB 64 V---LMVALYHLP---QLAGAVLLALVW-----WTLATVLLVLT-----PE 99  
QY 124 ARNYL--RDTSGVIFVLTWPLFGSFAAMLQNMNSIPGTYFILTFLMCLVIAVDGGYI 181  
DB 100 SVGYWGRWRRLGMGLILLPAQGLVLLKQWPLANGL-----IIAVVVLVWAGDIGAYF 154  
QY 182 AGVFFGSHPMAPLVSPKSWEGPAGSIVLG-SVTGALSVHFLLDHHWVG-----VILGA 236  
DB 155 SGKAFGRKRLAPRVSPKSGWEGYGLAASLAITLAVGLY-----RGWSIGALLALLGAA 210  
QY 237 LVVCAVL-GDLVESQFKDLGIKDKSNLPGHGLMDRLDGLMPLA-----AMVTW 285  
DB 211 LVVFSIVGDLTESMFKRSGINDSSNLLPGHGVLDRLDSLTAAIPVPAALLW 264  
RESULT 4  
CDS: ECOLI  
ID CDS: ECOLI STANDARD; PRT: 249 AA.  
AC P06466;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
synthase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
synthase) (CDS) (CTP-phosphatidate cytidyltransferase) (CDP-DAG  
synthase) (CDP-DG synthetase).  
GN CDS OR CDS OR B0175 OR Z0186 OR ECS0177 OR SF0165 OR S0168.  
OS Escherichia coli,  
OS Escherichia coli O157:H7, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 83334, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli;  
RX MEDLINE=86008268; PubMed=2995358;  
RA Icho T., Sparrow C.P., Raetz C.R.H.;  
RT "Molecular cloning and sequencing of the gene for CDP-diglyceride  
synthetase of Escherichia coli.";  
RL J. Biol. Chem. 260:12078-12083(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=KL2 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1233-1238(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=KL2 / W3110;  
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,  
Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,  
Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;  
RT "Systematic sequencing of the Escherichia coli genome: analysis of the  
4.0 - 6.0 min (189,987 - 281,416bp) region.";  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli;  
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,

RA Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
RA Davis R.W.,  
RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=0157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Bourin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potancusis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.,  
RA "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
RT Nature 409:529-533 (2001).  
RN [6]  
RN SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=0157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,  
RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RT DNA Res. 8:111-22 (2001).  
RN [7]  
RN SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=2272406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.,  
RA "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity  
RT through comparison with genomes of *Escherichia coli* K12 and O157.";  
RT Nucleic Acids Res. 30:4432-4441 (2002).  
RN [8]  
RN SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.,  
RA "Complete genome sequence and comparative genomics of *Shigella*  
RT flexneri serotype 2a strain 2457T.";  
RT Infect. Immun. 71:2775-2786 (2003).  
RN [9]  
RN CHARACTERIZATION.  
RC SPECIES=E.coli;  
RX MEDLINE=86008269; PubMed=2995359;  
RA Sparrow C.P., Raetz C.R.H.,  
RA "Purification and properties of the membrane-bound CDP-diglyceride  
RT synthetase from *Escherichia coli*.";  
RL J. Biol. Chem. 260:12084-12091 (1985).  
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
CC diacylglycerol.  
CC -!- PATHWAY: Phospholipid biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -!- SIMILARITY: Belongs to the CDS family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL; M11330; AAA23545.1; -;  
DR EMBL; AE000127; AAC73286.1; -;

DR EMBL; D83536; BAA77850.1; -;  
DR EMBL; U70214; AAB08604.1; -;  
DR EMBL; AE005193; AAG54477.1; -;  
DR EMBL; AP02550; BAB33600.1; -;  
DR EMBL; AE015053; AAN41827.1; ALT\_INIT.  
DR EMBL; AE016978; AAP15708.1; -;  
DR PIR; A23898; SYECDG.  
DR PIR; A85502; A85502.  
DR PIR; A99651; A99651.  
DR EcoGene; EGI0139; cdaA.  
DR InterPro; IPR000374; PC trans.  
DR Pfam; PF01148; CTP trans\_1; 1.  
DR PROSITE; PS01315; CDS; 1.  
KW Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;  
KW Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 20 40 POTENTIAL.  
FT TRANSMEM 57 77 POTENTIAL.  
FT TRANSMEM 85 105 POTENTIAL.  
FT TRANSMEM 115 135 POTENTIAL.  
FT TRANSMEM 154 174 POTENTIAL.  
FT TRANSMEM 177 197 POTENTIAL.  
FT TRANSMEM 228 248 POTENTIAL.  
SQ SEQUENCE 249 AA; 27571 MW; F628EC847E193647 CRC64;  
Query Match 18.2%; Score 285; DB 1; Length 249;  
Best Local Similarity 31.0%; Pred. No. 18-14; Indels 64; Gaps 12;  
Matches 85; Conservative 38; Mismatches 87;  
QY 47 WG-----WYLVAGEMAATMEVGRKEGGYHLPIMIGQAIWLS--- 91  
DB 7 WQQLSGFTTRSORVMLAVCGLLALLMLFL---LPEYHRNIHQPLVEIS---LWASLW 59  
QY 92 WPGGTGILASVAIVLVLMYFRIFNGTEKARNYLRDTSVGIF-VLTWPLF-GSEAA 149  
DB 60 W-----IVALLVL-----FYPGSAATWRN--SKTLRIFGVITVFFWGMAL 102  
QY 150 MSLMQNNSIPGTFTPLMCLVIASDVGGYIAGVFFGSHPNAPLVSPKSKWEGFAGSIV 209  
DB 103 RAWHYDENHYSGAINLLYVILWVGADSGAYWFGKLFGRKLPKVPKTKWQGFIGGLA 162  
QY 210 LGSVTGALSHVFLDHHWNGV-----ILGALV--VCATLGLDVSQFKREDLGI 257  
DB 163 TRAVIS-----WGYGWNLVDVAPVTLTLCISVAALASVLGDLTSMFKREAGI 211  
QY 258 KQMSNLLPGHGLMDRLDGMPLPAAVMTWLISVI 291  
DB 212 KUSCHLIPGHGILDRIDSLTRAAPVFACLLLV 245  
RESULT 5  
CDS\_CDS\_HAEIN STANDARD; PRT; 288 AA.  
AC P44937; O32623; O32627;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG  
DE synthase) (CDP-DG synthetase).  
DE CDS OR CDS OR H10919.  
GN Haemophilus influenzae.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RX [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=rd/kw20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness B.F.,  
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,



QY 134 GIVLTIWIPGSPAAFLSLMONSIPCTYFILTFMLCVIASDVGGYIAGVFGSHPMAP 193  
 Db 108 --FLFYALP-----LYLKIKEI-STYLLIFIGIWINDFYAYIGNFKTKLFP 159  
 QY 194 LVSPKKSWEFGAGSIVLGSVTVGSAVHF-LDDHHWMMGVILGCALVVCATLGDVIVESQFK 252  
 Db 160 KISPXKTEVEGLGGVLFGLFALTPYGIILN-----SFLGLTEVLTGVAGDYFKSFIK 214  
 QY 253 RDLGKDKSNLLPGHGLMDRLDGLMLPAAMVTWIL 288  
 Db 215 RQVGKIDFSNVFGEHGGTDFDALVFSAPVFLIM 250  
 RESULT 7  
 CDSA\_BRUBAB  
 ID CDSA\_BRUBAB STANDARD; PRT; 270 AA.  
 AC Q59173;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
 DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
 DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG  
 DE synthase) (CDP-DG synthetase).  
 GN CDSA.  
 OS Brucella abortus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S2308;  
 RA Bearden S.W., Ficht T.A.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
 CC diacylglycerol.  
 CC -!- PATHWAY: Phospholipid biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (By similarity).  
 CC -!- SIMILARITY: Belongs to the CDS family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U51683; AAA96785.1; -;  
 CC InterPro; IPR000374; PC\_Trans.  
 CC Pfam; PF01148; CTP\_transf\_1; 1.  
 CC PROSITE; PS01315; CDS; 1.  
 KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;  
 KW Transmembrane; Inner membrane.  
 FT TRANSMEM 19 39 POTENTIAL.  
 FT TRANSMEM 53 72 POTENTIAL.  
 FT TRANSMEM 76 96 POTENTIAL.  
 FT TRANSMEM 101 121 POTENTIAL.  
 FT TRANSMEM 126 146 POTENTIAL.  
 FT TRANSMEM 248 268 POTENTIAL.  
 FT TRANSMEM 270 AA; 28448 MW; 47253043415FF06 CRC64;  
 SQ SEQUENCE 270 AA; 28448 MW; 47253043415FF06 CRC64;  
 Query Match 15.6%; Score 244; DB 1; Length 270;  
 Best Local Similarity 27.1%; Pred. No. 1.4e-11;  
 Matches 73; Conservative 35; Mismatches 97; Indels 64; Gaps 9;  
 QY 81 IIGGQAIWLSWPPFGMGILASGVATVLMVFRIFYNGTEKEARNYLRTSVG-IFVLT 139  
 Db 12 IVLGITITLWLTW-VGGVGFTLSIAIGLAV-----FYEWTELSA---ITQTAFSRLFGWA 62

QY 140 WI-----PLFGSF-----AAMLSIMQN 156  
 Db 63 WLIVTIGILLDRGALLTIGFLVAGCAILLVTQWKGWPAAGLFYAGFSALSLLRG 122  
 QY 157 NSIPGTYFILTFLMGLVI-ASDVGGYIAGVFGSHPMAPLVSPKKSWEFGAGSIVLGSVTVG 215  
 Db 123 DE-PFGFTTIVFLFAVWVSTDTATFNGRAGLGPKLAPRFSPKNTWSAIGGAAAAGVAGG 181  
 QY 216 ALSVHFLLDHHWMMGVILGCALVVCATLGDVIVESQFKRDLGKDKSNLLPGHGLMDRLD 275  
 Db 182 LLVASLVAAPGCGMGVPLALLLSIVSQIGDLAESVVKRQFGAKDSRLLPGHGGVLDVDRD 241  
 QY 276 GMLPAAMVTWL-----ILSVISSY 295  
 Db 242 GLVAAAALLYLFGAIPAEPPVLSAIFFSF 270  
 RESULT 8  
 CDSA\_BRUSU  
 ID CDSA\_BRUSU STANDARD; PRT; 270 AA.  
 AC Q8G0E0;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
 DE synthetase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG  
 DE synthase) (CDP-DG synthetase).  
 GN CDSA OR BR1157.  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29461;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eissen J.A., Heidelberg J.F.,  
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beaman M.J.,  
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,  
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;  
 RT "The Brucella suis genome reveals fundamental similarities between  
 RT animal and plant pathogens and symbionts."  
 RL Proc Natl Acad Sci U.S.A. 99:13148-13153(2002).  
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
 CC diacylglycerol.  
 CC -!- PATHWAY: Phospholipid biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (By similarity).  
 CC -!- SIMILARITY: Belongs to the CDS family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB014416; AAN30077.1; -;  
 CC TIGR; BR1157; -;  
 CC InterPro; IPR000374; PC\_trans.  
 CC Pfam; PF01148; CTP\_transf\_1; 1.  
 CC PROSITE; PS01315; CDS; 1.  
 KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;  
 KW Transmembrane; Inner membrane.  
 FT TRANSMEM 19 39 POTENTIAL.  
 FT TRANSMEM 53 73 POTENTIAL.  
 FT TRANSMEM 76 96 POTENTIAL.  
 FT TRANSMEM 101 121 POTENTIAL.  
 FT TRANSMEM 126 146 POTENTIAL.





— — — — —

QY 142 PLFGSFAAMLSMQNSIPGTYFILTFLMCLVIAISDVGGYIACVFFGSHPMAPLVSPKSW 201  
 Db 117 -LFGVYKDF-----GAVALIWLIVVAVGDAFFGCKLGTPTT-ATSPNKL 164  
 QY 202 EGFAGSIVLGSVTGALS VHFLLDHHWGMVILGALVVCATGLDVLVESQFKRDILGIXDMS 261  
 Db 165 EGALIGVVLASVLSGFSVGMKLSGFLMALFLSLIALMAVFGDLYESYLKRRKVGVDG 224  
 QY 262 NLLPCHGGLMDRLDGMPLPAAMVTWLIL 288  
 Db 225 KILPGHGGVLDRLDLSMLFGALSLHLVL 251

## RESULT 12

YNBB\_ECOLI STANDARD; PRT; 298 AA.

AC P76091; P78229;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein ynbB.  
 GN YNB5 OR B1409.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474 (1997).  
 RN [2]

SEQUENCE OF 187-298 FROM N.A.

RC STRAIN=K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,  
 RA Sampaio G., Seki Y., Sivasubram S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377 (1996).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the CDS family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF000238; AAC74491.1; -;  
 CC EMBL; D90779; BAAL5023.1; -;  
 CC EMBL; D90780; BAAL5027.1; -;  
 CC PIR; D64892; D64892.  
 CC EcoGene; EG13749; ynbB.  
 CC InterPro; IPR000374; PC trans.  
 CC Pfam; PF01148; CTP transf.1; 1.  
 CC PROSITE; PS01315; CDS; 1.  
 CC Hypothetical protein; Transmembrane; Transferase; Complete proteome.

KW Hypothetical protein; Transmembrane; Transferase; Complete proteome.  
 FT TRANSMEM 5 25 POTENTIAL.  
 FT TRANSMEM 52 72 POTENTIAL.  
 FT TRANSMEM 105 125 POTENTIAL.

FT TRANSMEM 138 158 POTENTIAL.  
 FT TRANSMEM 163 183 POTENTIAL.  
 FT TRANSMEM 208 228 POTENTIAL.  
 FT TRANSMEM 236 256 POTENTIAL.  
 FT TRANSMEM 273 293 POTENTIAL.  
 FT TRANSMEM 187 197 W -> M (IN REF. 2).  
 FT CONFLICT 190 190 S -> I (IN REF. 2).  
 SQ SEQUENCE 298 AA; 33053 MW, 8083BA7E3C07FADF CRC64;

Query Match 14.7%; Score 230.5; DB 1; Length 298;

Best Local Similarity 26.1%; Pred. NO. 1.6e-10;  
 Matches 7%; Conservative 47; Mismatches 121; Indels 47; Gaps 11;

QY 12 RMPEKPNAGRDLCRAIAVIGIGLVILGIVLSPGW----YILVAGFMAAAATWEVGR 67  
 Db 26 RLPERGG-----EVLRIRTWGVICFMSWISGRWMTLTFALISFLA----- 71  
 QY 68 LKEG----GYHLPIMIGQAIWLSPFGTGMILASFVATVILMY-----PRIFYN 118  
 Db 72 LKEYCTLSVHPR-----WLYGIPLNVLLIGFNCFFELFLFIPLAGFLILAT 120  
 QY 119 G--TEKEARNYLDTSVGIF---VLTWIFLGSFAAMLSMQNSIPGTYFILTFLMCLVI 173  
 Db 121 GOVLVGDPSGFLHTVS-AIFWGMIMTFAL--SHAALLMLPTTNIQGGALLVFLILALT 177  
 QY 174 AS-DVGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIVLGSVTCALSVHFLDHHWGMVI 232  
 Db 178 ESNDIAQVLWGKSCGRKVKVPKVTLEGLMGVITIMIASLIIGPLLTPLNTLQALL 237  
 QY 233 LGCALVVCATGLDVLVESQFKRDILGKIDMSNLLPHGGLMDRLDGMPLPAAVY 283  
 Db 238 AGLLIGISGFGDGVVMSAIKRDIGVSKLPGHGLDRLDLSIFTAPV 288

## RESULT 13

CDSA\_SINY3 STANDARD; PRT; 293 AA.

AC P73548;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride  
 DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol  
 DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG  
 GN CDSA OR SLR1369.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirose M., Sugita M., Saito M., Saito T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136 (1996).  
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-  
 CC diacylglycerol.  
 CC -!- PATHWAY: Phospholipid biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the CDS family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

or send an email to license@isb-sib.ch).
-----
CC EMBL; D90907; BAA17588.1; -.
CC PIR; S77254; S77254.
CC InterPro; IPR000374; PC trans.
CC Pfam; PF01148; CTP trans1; 1.
CC PROSITE; PS01315; CDS; 1.
KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Complete proteome.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 97 117 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 273 293 POTENTIAL.
SQ SEQUENCE 293 AA; 31875 MW; A/E3ABC9C17475D0 CRC64;

Query Match 14.3%; Score 224; DB 1; Length 293;
Best Local Similarity 27.9%; Pred. No. 4.7e-10;
Matches 87; Conservative 49; Mismatches 112; Indels 64; Gaps 16;

22 RDLKAAIAGVIGLVVLVILGVLSPGWYIIVAGHMAAATWEGSRLE----- 70
Db 5 RIISAVIGIALAFSLIL-----GGWY-----FSAAILVIVLGLREYFQWVRAGIA 52
QY 71 --GGVHLPPIIMIGQAIWLWSP-----FGTIGILASF-----VATV--LVLMYF 113
Db 53 PRAKTTWLSLMLLSAIVT-----PHLDAFPPLTGALICVYLLFQPKMATIADISTLL 108
QY 114 RIFYNG--TEKEARNYLRDTSVIGFVLTWIPFGSFAAMLSLQNNISIPGYFILTFLMC 171
Db 109 GLFYGYGLPSYVWRLRGLDGVANPMGL--HLPLNGFWPESWAHPNFTGLLVTLIAF--AC 166
QY 172 VTASDVGVYGVVFGVSHPMAPLVSPKSWE-----GFGASIVLGSVGTALSVEHFLDHHW 227
Db 167 IWAADIGAYTMKWLGRTRLSL--ISPKKTVEGSLWGVGSLVG--VLGA-----WYLOWPY 220
QY 228 W--MGVILGALVAVCATGLDVLVESQFRLDGIKMSNLLPGHGLMDRLDGLMPLAAMVTW 285
Db 221 WEITGALLGLLIGIVSLGLDLTESMMKRDAGVDSGQLIPGHGGILDRTSDSYVFTAPLVY 280
QY 286 ----LILSVISS 293
Db 281 YFVILLPLVNN 292

RESULT 14
CDS: R1CPR STANDARD; PRT; 228 AA.
AC Q9ZDAB;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS: OR RP424.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichert-Ponten T., Almqvist U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."

```

```

RL Nature 396:133-140(1998).
CC -I- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
CC diacylglycerol.
CC -I- PATHWAY: Phospholipid biosynthesis.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -I- SIMILARITY: Belongs to the CDS family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ235271; CAA14881.1; -.
CC PIR; G71700; G71700.
CC InterPro; IPR000374; PC trans.
CC Pfam; PF01148; CTP trans1; 1.
CC PROSITE; PS01315; CDS; 1.
KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
SQ SEQUENCE 228 AA; 25814 MW; 48DE66AB6EA19174 CRC64;

Query Match 13.9%; Score 216.5; DB 1; Length 228;
Best Local Similarity 29.1%; Pred. No. 1.4e-09;
Matches 67; Conservative 41; Mismatches 77; Indels 45; Gaps 11;

QY 79 IMIIGQAIWLWSPFGTIGILASFVATV--VLMYFRI-----FYNGTEKAR 125
Db 19 LRIISGIALVSL-----FVAILCLKTLFYILMLVGLGMLSEWYNTYPSI- 65
QY 126 NYLRDTSVIGFVLTWIPFGSFAAMLSLQNNISIPGYFILTFLMCVIA--DVGYIAGV 184
Db 66 NYL---LIGLII---IPIPIISLILFSLMBESNRL-----VIMLYFCILWSDVTFAMIGK 114
QY 185 FFGSHPMAPLVSPKSWEG--FAGSIVLG--SVTGALSVEHFLDHHWVG---VILGAL 237
Db 115 TFKGIKLAKIPSKTWTGLITGTVSAGLVSVLSIPNYHIEHYFYSKIVLFIISLIL 174
QY 238 VVCATGLDVLVESQFRLDGIKMSNLLPGHGLMDRLDGLMPLAAMVTWLI 287
Db 175 ALIAQSSDLFISYFKRKFNKIDSGHIIPGHGGVLDKDFDSIILTAPVFCI 224

RESULT 15
CDS: R1CPR STANDARD; PRT; 270 AA.
AC Q9X1B7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS: OR TM1397.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

```

Search completed: August 11, 2004, 13:35:09  
Job time : 15 secs

Search completed: August 11, 2004, 13:35:09  
Job time : 15 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:32:07 ; Search time 39 Seconds  
(without alignments)  
2402.792 Million cell updates/sec

Title: US-09-853-641-2  
Perfect score: 1563  
Sequence: 1 MNEPEQHRSRMFPKPKNA.....LPAAVVTWLLSVISSYPS 297

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*\*  
1: sp\_archaea:\*\*  
2: sp\_bacteria:\*\*  
3: sp\_fungi:\*\*  
4: sp\_human:\*\*  
5: sp\_invertebrate:\*\*  
6: sp\_mammal:\*\*  
7: sp\_mhc:\*\*  
8: sp\_organelle:\*\*  
9: sp\_phage:\*\*  
10: sp\_plant:\*\*  
11: sp\_rodent:\*\*  
12: sp\_virus:\*\*  
13: sp\_vertebrate:\*\*  
14: sp\_unclassified:\*\*  
15: sp\_virus:\*\*  
16: sp\_bacteriap:\*\*  
17: sp\_archaeap:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1495	95.6	285	16	Q8NP04
2	1366	87.4	297	16	Q8FP76
3	551	35.3	392	16	Q82JY1
4	539	34.5	391	16	Q86769
5	350.5	22.4	328	16	Q8G482
6	322.5	20.6	280	16	Q83G74
7	322	20.6	269	16	Q83NN2
8	310	19.8	271	16	Q885N8
9	310	19.8	288	16	Q7VRE0
10	299	19.1	271	16	Q88MH5
11	299	19.1	282	16	Q8ZH60
12	288.5	19.1	289	16	Q7VM21
13	298	19.1	280	16	Q9KPV7
14	298	19.1	285	16	Q8XG59
15	296	18.9	285	16	Q8FL09
16	291	18.6	272	16	Q83BV6

17	286.5	18.3	265	16	Q9JX34
18	282	18.0	275	16	Q8PML0
19	280.5	17.9	265	16	Q9K1G7
20	280.5	17.9	289	16	Q9CJL3
21	279	17.9	275	16	Q8PAV8
22	278	17.8	280	16	Q8A0L5
23	278	17.8	280	16	Q89XP8
24	276.5	17.7	285	16	Q7TU77
25	274.5	17.6	273	16	Q8KFT9
26	272	17.4	282	16	Q8DBF6
27	271	17.3	280	16	Q87ME2
28	268	17.1	271	16	Q8XZ16
29	265.5	17.0	294	16	Q8YQF6
30	265	17.0	348	16	Q7USC6
31	259.5	16.6	284	16	Q9RU83
32	257.5	16.5	296	16	Q7VB11
33	256.5	16.4	277	16	Q92Q50
34	254	16.3	298	16	Q7U7N9
35	253	16.2	265	16	Q895K7
36	253	16.2	399	10	Q9M001
37	252.5	16.2	391	10	Q94A03
38	252.5	16.2	430	10	Q7XJM7
39	252	16.1	309	16	Q87TX6
40	251	16.1	310	16	Q8PQ11
41	250.5	16.0	264	16	Q8XJRO
42	250	16.0	304	16	Q7TUM2
43	249	15.9	241	16	Q9PMW2
44	248.5	15.9	278	16	Q8YHH2
45	247.5	15.8	260	16	Q8CST9

## ALIGNMENTS

## RESULT 1

Q8NP04 PRELIMINARY; PRT; 285 AA.

AC Q8NP04; 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE CDP-diglyceride synthetase (EC 2.7.7.41).

GN CGL2022.

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

OX NCBI\_TaxID:1718;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RA Nakagawa S.;

RL "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

DR EMBL; AP005280; BAB99415.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.

DR InterPro; IPR000374; PC trans.

DR Pfam; PF01148; CTP transf\_1, 1.

DR PROSITE; PS01315; CDS; 1.

KW Transferase; Nucleotidyltransferase; Complete proteome.

SQ SEQUENCE 285 AA; 30733 MW; E41025439D7002B4 CRC64;

Query Match 95.6%; Score 1495; DB 16; Length 285;  
Best Local Similarity 100.0%; Pred. No. 1.5e-107;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 MPKPKNNAGRDLLKAAIAVGIGLVLLGIVLSPGWVILVAGFMAATWVGSRLEKGG 72

Db 1 MPKPKNNAGRDLLKAAIAVGIGLVLLGIVLSPGWVILVAGFMAATWVGSRLEKGG 60

QY 73 YHLPLPIMIGQAIWLSPFGTGMGLASFVATVILVLMYFRIFYNGTEKEARNYLRTS 132

```

Db 61 YHPLPIMIGQALIWSPFGTGMGILASPVATVLMVFRIFYNGTEKEARNYLRDTS 120
QY 133 VGIFVLTWIPFGSFAAMLSMNNISPGTYFILTMLCVIASDVGGYIAGVFGSHPMA 192
Db 121 VGIFVLTWIPFGSFAAMLSMNNISPGTYFILTMLCVIASDVGGYIAGVFGSHPMA 180
QY 193 PLVSPKXWEGFAGSIVLGSVTGALSVDHLDHMMWGVILGCAVVCATIGDLVVSQFK 252
Db 181 PLVSPKXWEGFAGSIVLGSVTGALSVDHLDHMMWGVILGCAVVCATIGDLVVSQFK 240
QY 253 RDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVSSYPS 297
Db 241 RDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVSSYPS 285

RESULT 2
Q8FP76 PRELIMINARY; PRT; 297 AA.
AC Q8FP76;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative phosphatidate cytidyltransferase.
GN CDS OR SAV2623.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR ENBL: AP005031; BAC70334.1;
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR000374; PC trans.
DR Pfam: PF01148; CTP trans_1; 1.
DR PROSITE: PS01315; CDS; 1.
DR Nucleotide transferase; Transferase; Complete proteome.
KW NUCLEOTIDYLTRANSFERASE; TRANSFERASE;
SQ SEQUENCE 297 AA; 32053 MW; 634F0906B4E06915 CRC64;

Query Match 87.4%; Score 1366; DB 16; Length 297;
Best Local Similarity 83.7%; Pred. No. 1.3e-97;
Matches 247; Conservative 30; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNEPEQHRSRMMPKPKNAGRDLLKAAIAGVGLGVLLGIVLSPGWYILVAGFMAA 60
Db 1 MSEPHEHRTVMRKPKNAGRDLLKAAIAGVGLGVLLGIVLSPGWYILVAGFMAA 60
QY 61 TWEVSGRLKEGYHPLPMTIIGQAIILWSPFGTGMGILASPVATVLMVFRIFYNGT 120
Db 61 TWEVCGRLKEAGYHPLPMTIIGQAVVWCSWFFETSGILASYATVLMVFRIFYNGP 120
QY 121 EKEARNYLRDTSVGIFVLTWIPFGSFAAMLSMNNISPGTYFILTMLCVIASDVGGY 180
Db 121 EKEARNYLRDTSVGIFVLTWIPFGSFAAMLSMNNISPGTYFILTMLCVIASDVGGY 180
QY 181 IAGVFGSHPMAPLVSPKXWEGFAGSIVLGSVTGALSVDHLDHMMWGVILGCAVVC 240
Db 181 AAGVFGSHPMAPLVSPKXWEGFAGSIVLGSVTGALSVDHLDHMMWGVILGCAVVC 240
QY 241 ATIGDLVVSQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVSSY 295
Db 241 ATIGDLVVSQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVSSY 295

```

```

RESULT 3
Q82JY1 PRELIMINARY; PRT; 392 AA.
AC Q82JY1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative phosphatidate cytidyltransferase.
GN CDS OR SAV2623.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR ENBL: AP005031; BAC70334.1;
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR000374; PC trans.
DR Pfam: PF01148; CTP trans_1; 1.
DR PROSITE: PS01315; CDS; 1.
DR Nucleotide transferase; Transferase; Complete proteome.
KW NUCLEOTIDYLTRANSFERASE; TRANSFERASE;
SQ SEQUENCE 392 AA; 40876 MW; C6F6A4254E7C8AC6 CRC64;

Query Match 35.3%; Score 551; DB 16; Length 392;
Best Local Similarity 39.1%; Pred. No. 1.3e-34;
Matches 117; Conservative 52; Mismatches 114; Indels 16; Gaps 5;

QY 3 EPEQHRSRMMPKPKNAGRDLLKAAIAGVGLGVLLGIVLSPGWYILVAGFMA 57
Db 102 QPKPEPMPDAPQAPAPQPKSAGRDLLGAAIAGVGLGVLLGIVLSPGWYILVAV 161
QY 58 AATWVSGRLKE-GGYHPLPMTIIGQAIILWSPFGTGMGILASPVATVLMVFRIF 116
Db 162 VGLWELTSLERBERKIKAPLPLVAGGAAYVAGVYRGAEGAVAMALTALAVLWRM- 220
QY 117 YNGTEKEARNYLRDTSVGIFVLTWIPFGSFAAMLSMNNISPGTYFILTMLCVIASD 176
Db 221 ---TEPPEGYLRDVTAGVFAAFYVFFLATFVAMMLTADD---GFRRLTFLTLTVSD 272
QY 177 VGGYIAGVFGSHPMAPLVSPKXWEGFAGSIVLGSVTGALSVDHLDH-HMMWGVILGC 235
Db 273 TGAATAGWRFGRHRLAPRISPKTREGVLGVASFANVAGALCMFEFLDDSGWQGLVGF 332
QY 236 ALYVVCATLGDVVSQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVSS 294
Db 333 AVAASATLGDGSSMIKRDGLKIDMSNLLPGHGLMDRLDLSLLPTAPVWVLLLVFVGS 391

RESULT 4
Q86769 PRELIMINARY; PRT; 391 AA.
ID Q86769
AC Q86769;

```



Dd      278 VQWREGRKKLAPRISPKGTREGLLGATAFANVAGALCMQLDIDGGAWQGILLGLVVAVS    333

Qy      241 ATFLGDLVESQFKRLGIKMSNLLPGRGGLMDRLDGMLPAAMVTWLILSVISS    294  
       |||||                |||||                |||||                ||  
Db      338 ATLGLGESMIKRDLGIKMGTLLPGHGGIMDRLDLSLLPTAPVVLLLIVFVS    391  
       |||||                |||||                |||||                ||

RESULT 5

QB8q482 PRELIMINARY; PRT; 328 AA.

ID QSG482;  
AC QSG482;  
DT 01-WAR-2003 (TremBLrel. 23, Created)  
DT 01-WAR-2003 (TremBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Probable phosphatidate cytidyltransferase.  
CDS OR BLI507.  
OS Bifidobacterium longum.  
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
CC Bifidobacteriaceae; Bifidobacterium.  
NCBI\_TaxID=216816;  
[ ]\_TaxID=  
RN [ ]\_  
XP SEQUENCE FROM N.A.  
RP STRAIN=NCC 2705;  
RC MEDLINE=22294377; PubMed=12381787;  
RX Schell M.A., Karimiantzou M., Snel B., Vilanova D., Berger B.,  
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,  
RR Primore R.D., Arigoni F.;  
RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
to the human gastrointestinal tract.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
RD EMBL; AB014782; AAN25302.1; -;  
RG GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.  
DD DR; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.  
DR InterPro; IPR000374; PC\_trans.  
DR Pfam; PF01448; Clp\_transf\_l; l.  
KW Nucleosidyldiphosphate transferase; Transferase; Complete proteome.  
SQ SEQUENCE 328 AA; 35279 MW; B4SCBA105FC37839 CRC64;

Query Match 22.4%; Score 350.5; DB 16; Length 328;  
Best Local Similarity 31.0%; Pred No. 3.le-19;  
Matches 104; Conservative 52; Mismatches 119; Indels 61; Gaps 12

QY 1 MNPEQHRSMR--MPKPQNAGRDKAAIAVGIGVLGVLLIGIVLSPGWIKLVAGFWA 58  
Db 1 MERNEQPHEEAETLDIQNKKTGNRMPOAITGAALVILIACLLSLVDLFVLLVVFFM 60  
QY 59 RATEVGSRLKEGYHLPLPIIMIIGAAILWSWPFG-----TMG-----I 99  
Db 61 LALWEIRVDFPATVGHLPV-----FMULWCSSFTLIATYSPPHLTNGLSIIVIS 112  
QY 100 LASFYATVLYLMYFRIF-----YNGTEKEAR-----NYLRDTSYGIFVLT-- 139  
Db 113 LVIAIASAKISVGNRKLSLAVAGKLSDNTASARLESSFNHGEGSQHHSRSHSAVSVLTV 172  
QY 140 WTLPGSPAAMLSQLMQNNISPICYFIITFMLCVIADVGVIAGVFSGHPMAPLVSPKK 199  
Db 173 YIPLAS-CVIISTTFNFHGPFAHAMLVFPAL--SDTGGLFAGAWLGKHKSPPRISSPK 229  
QY 200 SMEGFAGSIGVSGVTGALSVEHFLLDH-----HWMMGVILTGCALVVCAITGLDVESQEK 252  
Db 230 SVBGLVGS-MLFAMAGAFAV-FACTYDASKWAIEWVPVIGILICAVGTFGDLCASMVK 287  
QY 253 RDLGTKMSNLDPHQGLMDRLDGMLPAAMVTWLIL 286  
Db 288 RDIGIKDMGHLLKGHGMVDRVDSLIMSAPTFCALL 323

RESULT 6

QB8g74 PRELIMINARY; PRT; 280 AA.

ID QB8g74  
AC QB8g74;

```

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41).
GN CDS A OR TW445.
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RA Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,
RA Claverie J.-M.;
RA "Tropheryma whipplei illustrates the diversity of gene loss patterns
RT in small genome bacterial pathogens.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016851; AA044542.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000374; PC_trans.
DR Pfam; PF01148; CTP_transf_1.
DR PROSITE; PS01315; CDS; 1.
RW Complete proteome.
SQ SEQUENCE 280 AA; 31290 MW; 6E64036636D5D6A9 CRC64;

Query Match 20.6%; Score 322.5; DB 16; Length 280;
Best Local Similarity 31.7%; Pred. No. 3.8e-17;
Matches 92; Conservative 47; Mismatches 114; Indels 37; Gaps 10;

QY 1 MNEPEO-HRSMRMPKPNAGRDKAATVAGIGLVLLG--IVLSPWGYILVAGFM 57
Db 1 MNEAMLRVFFMLKQVQARTGRDLVLAIGVGLILGGLLISLFVFLHFLVLCILVGT 60
QY 58 AATWVGSRLEGGVHLPETM-----IIGQAIWLSWPGTGMGLASFVATVLMY 112
Db 61 CLA--EIFTATHTRGIFIVRFLTSTVLPVLAFFASW---FLPSIAFFMLTLLAQF 115
QY 113 FRIFYNGTEKARNYLRDTSVGFVLTWLPFGSFAAMLSLMQNNISIPGTIFILTFMLCV 172
Db 116 F-----LRARFSLH-----SLIFLYIPITVSLFAIAAHT---GRLWVFFMLTV 159
QY 173 IASDVGYYAGVFFGSHMPAPVSPKSWEGFAG----SIVLGSVTGALSVEFLDHHW 228
Db 160 IASDTSYVFGTLFGRHLLAPRISPKTWEGLMGGFFSFLFGTLTGILLHKSLL---L 215
QY 229 MGVLGALVVCATLGLDVSQFKDLGKDMNSLLPGHGGMLDRDGLM 278
Db 216 FSATAASILFLFALLGLDLAESYIKRLGVKDMGSLPGHGGMLDRDVSML 265

RESULT 7
ID Q83NN2 PRELIMINARY; PRT; 269 AA.
AC Q83NN2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative integral membrane phospholipid biosynthetic
DE nucleotidyltransferase.
GN TW322.
OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=218496;
RN [1]
RP SEQUENCE FROM N.A.
RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
RA Dover L.G., Norbertczak H.T., Besta G.S., Quail M.A., Harris D.E.,
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
RA Barrell B.G., Parkhill J., Rellman D.A.;

```

```

RT "Sequencing and analysis of the genome of the Whipple's disease
RT bacterium Tropheryma whipplei.";
RL Lancet 361:637-644 (2003).
DR EMBL; BX251411; CAD66995.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000374; PC_trans.
DR Pfam; PF01148; CTP_transf_1.
DR PROSITE; PS01315; CDS; 1.
RW Complete proteome.
SQ SEQUENCE 269 AA; 23898 MW; B91FB3AC0FC604C6 CRC64;

Query Match 20.6%; Score 322; DB 16; Length 269;
Best Local Similarity 32.7%; Pred. No. 4e-17;
Matches 88; Conservative 44; Mismatches 101; Indels 36; Gaps 9;

QY 21 GRDLKAAIAVIGIGLVLLG--IVLSPWGYILVAGFAAATWVGSRLEKGYHLPLP 78
Db 11 GRDLVLAIGVGLILGGLLISLFVFLHFLVLCILVGVTCIA--EIFTATHTRGIFIVR 68
QY 79 IM-----IIGQAIWLSWPGTGMGLASFVATVLMYFRIFYNGTEKARNYLRDTSV 133
Db 69 FLTLSTVLPVLAFFASW---FLPSIAFFMLTLLAQF-----LRARFSLH---- 113
QY 134 GIFVLTWLPFGSFAAMLSLMQNNISIPGTIFILTFMLCVIASDVGGYIAGVFFGSHMP 193
Db 114 SLIFLYIPITVSLFAIAAHT---GRLWVFFMLTVIASDTSYVFGTLFGRHLLAP 169
QY 194 LVSPKSWEGFAG----SIVLGSVTGALSVEFLDHHWVGVLGALVVCATLGLDVS 249
Db 170 RISFNKTWELGCGFFSSUFFGTLTGILLHKSLL-----LFSATAASILFLFALLGLAES 225
QY 250 QFKRDLGKDMNSLLPGHGGMLDRDGLM 278
Db 226 YIKRLGVKDMGSLPGHGGMLDRDVSML 254

RESULT 8
ID Q886N8 PRELIMINARY; PRT; 271 AA.
AC Q886N8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatidate cytidyltransferase.
GN CDS A-1 OR PSPT01539.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Koonin J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
RA White O., Fraser C., Collier A.;
RA "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016861; AAC55059.1; -.
DR TIGR; PSPT01539; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000374; PC_trans.
DR Pfam; PF01148; CTP_transf_1.
DR PROSITE; PS01315; CDS; 1.
RW Complete proteome.
RW Nucleotidyltransferase; Complete proteome.
SQ SEQUENCE 271 AA; 23308 MW; D737D038686E86B CRC64;

```

Query Match 19.8%; Score 310; DB 16; Length 271;  
Best Local Similarity 33.3%; Pred. No. 3.4e-16;  
Matches 99; Conservative 44; Mismatches 84; Indels 70; Gaps 17;

QY 22 RDLKAAIAGVIGL-GVLVLG-----IVLSPGWYILVAGFMAAATWEVGS-----65  
DB 5 RIITALLILPVALCGFFLTGYFALFGVGVVVLGAWEARL-AGF-AAQSMRVGYAALV 62  
QY 66 SRLKEGGYHLP--LPIMITGGQAIWLSPFGTGMGILASFVATVLVLMYFRIFYNGTEKE 123  
DB 63 AVLLFFMYLLPGLPEWLVLA--AVIWS-----VATFLVLT-----PD 99  
QY 124 ARYLRLDTS---VGFVLTWIPFGSPAAMLSLNNISIP-GTYFILTMLCVIADVNG 178  
DB 100 SSSHWASACKLVIGLLIL---LPWGLVLIKQWPLGNWLLSVLVWVLRADIG 151  
QY 179 GYIAGVFFGSHMAPLVSPKKSWEFGAGSIVLG-SVTGALSVEHFLDHHW-----WMGVIL 233  
DB 152 AYFSGKAFGRKLAPKVPSPKKSWEFGVGLVSLGITAAGVW-----RDWTWVQFIAALL 207  
QY 234 GCALVV-CATGLDVSQFKRDLGDKMSNLPLPGHGLMDRLDGLMPA-----NMVTW 285  
DB 208 GAAVIVTISVIGLTESMFKRSQVKSNNLLPFGHGVLDRLDRLDLSLTAIPVFAVLLW 264

RESULT 9  
Q7VRE0 PRELIMINARY; PRT; 288 AA.  
AC Q7VRE0;  
DT 01-OCT-2003 (TRENBLrel. 25, Created)  
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Phosphatidate cytidyltransferase (EC 2.7.7.41).  
GN CDSA OR BFL277.  
OS Candidatus Blochmannia floridanus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.  
OX NCBI\_TaxID=203907;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22784745; PubMed=12886019;  
RA Gil R., Silva P.J., Zient E., Delmorte F., Gonzalez-Candelas F.,  
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,  
RA van Ham R.C.H.J., Gross R., Moya A.,  
RT "The genome sequence of Blochmannia floridanus: comparative analysis  
RT of reduced genomes."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393 (2003).  
DR EMBL; BX248585; CAD83348.1; -;  
KW Nucleotidyltransferase; Transferase; Complete proteome.  
SQ SEQUENCE 288 AA; 33466 MW; 3D9F9FCAALD65F5A CRC64;

Query Match 19.8%; Score 310; DB 16; Length 288;  
Best Local Similarity 27.2%; Pred. No. 3.6e-16;  
Matches 83; Conservative 55; Mismatches 105; Indels 62; Gaps 9;

QY 22 RDLKAAIAGVIGLVLGIVLSPGWYILVAGFMAAATWEVGS--RLKEGGYHLPPI 79  
DB 5 RVISTCILPVLISILFLPSVT---QFSCLLFAVCAISAWENGKLMHFSNIYLL----56  
QY 80 MIIGQAIWLSPFGTGMGILASFVATVLV--LMYFRIF-----116  
DB 57 -----WIVFVCSILCITLIMILHNYLFNCFILLSCFSPVIMWVFMLLV 104  
QY 117 ----YNGTEKEARNLRDTSVGFVLTWIPFGSPAAMLSLMQ-----NNSIPGYFILTIF 168  
DB 105 LEVPYSSCVWNSNVL-----FCFGLLIIFPPFFGVLTKFGYSNNDVNGKWLIV 158  
QY 169 MLCVLTASDVGGIAGVFGSHMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFLD--HH 226  
DB 159 LVLVINDVSAVLIQGMFTHKLLKYVSPKTKTWEGFICGLVSTVIAIISKYIMIDVRNF 218  
QY 227 WNWGVILGALVVCATLGLDVSQFKRDLGDKMSNLPLPGHGLMDRLDGLMPAAMVTL 286

Db 219 YIFVCFVCA-ILFSVIGLTESMFKRSQVKTGNLIPFGHGLDRMDSIFSAPVTV 277  
QY 287 ILSVI 291  
Db 278 LLFLL 282

RESULT 10  
Q88MH5 PRELIMINARY; PRT; 271 AA.  
AC Q88MH5;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Phosphatidate cytidyltransferase.  
GN CDSA OR PPI596.  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=160488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22423060; PubMed=12534463;  
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzar A.,  
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
RA Lauber J., Stjepandic D., Hehseisel J., Straetz M., Heim S.,  
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,  
RA Fraser C.M.;  
RT "Complete genome sequence and comparative analysis of the  
RT metabolically versatile pseudomonas putida KT2440."  
RL Environ. Microbiol. 4:799-808 (2002).  
DR EMBL; AF015779; RAN67217.1; -;  
DR TIGR; PFI596; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.  
DR InterPro; IPR000374; PC trans.  
DR Pfam; PF01148; CTP transf. 1.  
DR PROSITE; PS01315; CDS; 1.  
KW Nucleotidyltransferase; Transferase; Complete proteome.  
SQ SEQUENCE 271 AA; 29075 MW; 68DD533670B5EB3E CRC64;

Query Match 19.1%; Score 299; DB 16; Length 271;  
Best Local Similarity 34.0%; Pred. No. 2.4e-15;  
Matches 100; Conservative 39; Mismatches 91; Indels 64; Gaps 15;

QY 22 RDLKAAIAGVIGLVLG-----IVLSPGWYILVAGFMAAATWEVGSRLKE 70  
DB 5 RIITALLILPVALCGFFLLNGDFFALFIGVVTILGAWEARL-AGLMAQPLRIAYAAVVA 63  
QY 71 GS-----YHLPPLPIMIGQAIWLSPFGTGMGILASFVATVLVLMYFRIFYNGTEKARN 126  
DB 64 GALLMLHLPELAPNVLGAIVW-WG-----LATWLVLTYPR-----SSDLWASA 107  
QY 127 YLRDTSVGFVLTWIPFGSPAAMLSLNNISIPGTFTILTMCLCVIADVGGYIAGVFF 186  
DB 108 ACR-LIIGLLVL--LPWQOQ---LVLLKHWPL-GNWLLSVMLVWVAADIGAYFSGRAF 159  
QY 187 GSHMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFL-----LDHHWWMGV-----LGCA 236  
DB 160 GKRKLAPQVSPKSWEG-----VYGLAVSLITLGVISRDGFGQILLGLGAA 210  
QY 237 LVVCAT-LGLDVSQFKRDLGDKMSNLPLPGHGLMDRLDGLMPA-----AMVTW 285  
DB 211 LLWSSVVDLTESMFKRSQVKSNNLLPFGHGVLDRLDLSLTAIPVFAVLLW 264

RESULT 11

```

RESULT 12
Q7VM21 PRELIMINARY; PRT; 289 AA.
ID Q7VM21;
AC Q7VM21;
DT 01-OCT-2003 (TtEMBLrel. 25, Created)
DT 01-OCT-2003 (TtEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TtEMBLrel. 25, Last annotation update)
DE Phosphatidate cytidyltransferase.
GN CSDA OR HD1193.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
NCBI_TaxID=730;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=5000HP / ATCC 700724;
RC Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017154; AAP96043.1; -
KW Nucleotidyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 289 AA, 31998 MW; FBBFC9F33BE3FEB7 CRC64;

Query Match 19.1%; Score 398.5; DB 16; Length 289;
Best Local Similarity 30.0%; Pred. No. 2.8e-15;
Matches 90; Conservative 39; Mismatches 104; Indels 67; Caps 9

QY 24 LKAAIAVAGIGLGVLLGI-VLSPGWYIIVAGFMAATWEVGRLEKGGYHLPPLMII 82
Db 2 LKERILSAIMIIVLAGIFLLSLPFTLLAAVITAAIWEA---QFAGFKSL----- 53
QY 83 GQQAIIWSPFGTWGIIASVAIVIVMTFRIFYNGTEKEARNYLRDTSVGIFVLT-- 139
Db 54 -SRAII-----TFVTICLFILL--IFANTHYIRAGRFELTDQTIPLLFISCCW 97
QY 140 MWPLFG-----SFAAMLMLQMN-----NSIPGTYFIL 166
Db 98 WLLIAGFLVISYPKASLWSKSIATKFIQFECTLIPIFFIALIRFNNAFNOFGAYLL 157
QY 167 TMLCVIASDVGGIAGVFGGSHMAPLVSPKKSWEFGAGSVLGSVTVGALS VH----- 220
Db 158 YVCLLWCADSGAYFFGFRAGFKGLAAKVSPKSWEGALGGLITSCVMALLFKLTPNNI 217
QY 221 FLLDHHWMGVILCALWCATIGDIVESQFKDLGIKDMSNLLPGHGGMRLDGNLPA 280
Db 218 FGRFEFTPLPIIVSVATVSSIGDLAESAESFKQAQIKDSSNLIPOGGGILDRISLTA 277

RESULT 13
Q9KPV7 PRELIMINARY; PRT; 280 AA.
ID Q9KPV7
AC Q9KPV7;
DT 01-OCT-2000 (TtEMBLrel. 15, Created)
DT 01-OCT-2000 (TtEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TtEMBLrel. 25, Last annotation update)
DE Phosphatidate cytidyltransferase.
GN VC2255.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
NCBI_TaxID=666;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RC MEDLINE=20406833; Pubmed=10952301;
RC Heidelberg J.F., Eise J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RR Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vanathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann P.D., Nierman W.C., White O.

```



OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Escherichia.
OX	NCBI_TaxID=217992;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=O6:HL / CFT073 / ATCC 70928;
RX	MEDLINE=22388234; PubMed 12471157;
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA	Rasko D.J., Buckles E.J., Llou S.-R., Boutin A., Hackett J., Stroud D. P.,
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT	"Extensive mosaic structure revealed by the complete genome sequence
RT	of uropathogenic Escherichia coli";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR	EMBL; AE016755; AAN78704.1; -.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0004603; F:phosphatidate cytidylyltransferase activity; IEA.
DR	GO; GO:0016740; F:transferase activity; IEA.
DR	GO; GO:0009654; P:phospholipid biosynthesis; IEA.
DR	InterPro; IPR000374; PC_trans.
DR	Fram; PF01148; CTP_transf_1; 1.
DR	PROSITE; PS01315; CDS; 1.
KW	Nucleocydyltransferase; Transferase; Complete proteome.
SQ	SEQUENCE 285 AA; 31481 MW; 4C35CA58351F0522 CRC64;
Query Watch 18.9%; Score 296; DB 16; Length 285;	
Best Local Similarity 29.7%; Pred. No. 4.3e-15;	
Matches 86; Conservative 43; Mismatches 99; Indels 62; Gaps	
QY	30 VGIGLVGLVGLIGLVSPWG-----WVILVAGFMAAATWVGSRKKEGGVHLPL
DB	:  :
DB	26 VGFAVLIVLCMAAENWGQISGFTTRSORVWLAVLCGLLALMLFL--LPYHNIHQ
QY	78 PIMIIGGQAIWIS--WPGTWGIIASFAVTLVLMYFIYNGTEKEARNYLRTDTSVG
DB	:  :
DB	83 PLVIRIS----LWASLGGW-----IVALLVL-----FYPGSAAIWRN-SKTLRII
QY	135 IFVLITWLPFL-GSPAAMI.SLMQNNISPGYFILTLMCLVIASDVGVIAGVFGSHMPAR
DB	:  :
DB	123 FGVLITVFFWGMALRAMHYDENHYSGAIWLVYVMIILWGADSGAIFYNKFIFKHKHAP
QY	194 LVSPKSGWEGFAGSIVLGSVGTGALSVFHLLDHHWMVG-----ILGCAIV--VCA
DB	:  :
DB	183 KVSPEKTWQGFICGLATAAVIS-----WGYGWNVLNVAPVTLICISVAALAS
QY	242 TLGDLIVSFQKRLGKDKSNLLPBGGLMDRLDQWLPAAMVTLILSVI 291
DB	:  :
DB	232 VLGLDTESNFKRAGIKDSGHLIPGGGTLIRDSLTAAVVFACLLLV 281

Search completed: August 11, 2004, 13:36:00  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:33:48 ; Search time 19 Seconds  
(without alignments)  
806.995 Million cell updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNPEQHRSMRMPKPKNA.....LPAAVTVLILSVISSYPS 297

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	19.1	500	4	US-09-252-991A-21214
2	295	18.9	289	4	US-09-489-039A-12390
3	285	18.2	249	4	US-09-282-218A-17
4	282.5	18.1	292	4	US-09-543-681A-6495
5	262	16.8	285	4	US-09-328-352-7257
6	255	16.3	315	4	US-09-543-681A-7121
7	248	15.9	296	4	US-09-540-236-2118
8	247.5	15.8	261	4	US-09-134-001C-4966
9	246.5	15.8	346	4	US-09-252-991A-31107
10	219	14.0	192	4	US-09-134-000C-5752
11	199	12.7	446	3	US-08-672-814D-11
12	199	12.7	446	3	US-09-333-696-11
13	199	12.7	446	4	US-09-282-218A-19
14	198.5	12.7	445	4	US-09-282-218A-12
15	198.5	12.7	461	2	US-08-672-814D-2
16	198.5	12.7	461	3	US-09-333-696-2
17	195.5	12.5	461	4	US-09-282-218A-2
18	193.5	12.4	308	4	US-09-198-452A-606
19	142	9.1	457	4	US-09-282-218A-18
20	116.5	7.5	238	4	US-09-107-532A-5419
21	115.5	7.4	326	3	US-08-986-788-2
22	110	7.0	492	4	US-09-489-039A-8343
23	109.5	7.0	397	4	US-09-543-681A-5521
24	109.5	7.0	494	4	US-09-489-039A-11852
25	105.5	6.7	597	4	US-09-252-991A-24252
26	105.5	6.7	637	4	US-09-489-039A-10817
27	105	6.7	354	1	US-08-313-553-9

28 105 6.7 354 3 US-08-767-993-9 Sequence 9, Appli  
29 103 6.6 285 4 US-09-252-991A-30476 Sequence 30476, A  
30 102 6.5 867 4 US-09-107-532A-4393 Sequence 4393, Ap  
31 100 6.4 323 4 US-09-489-039A-10164 Sequence 10164, A  
32 100 6.4 705 4 US-09-198-452A-68 Sequence 68, Appl  
33 99.5 6.4 305 4 US-09-134-001C-5038 Sequence 5038, Ap  
34 99.5 6.4 332 4 US-09-252-991A-31791 Sequence 31791, A  
35 99 6.3 493 4 US-09-134-001C-3486 Sequence 3486, Ap  
36 99 6.3 1033 4 US-09-252-991A-20611 Sequence 20611, A  
37 98.5 6.3 211 4 US-09-198-452A-1166 Sequence 1166, Ap  
38 98 6.3 256 4 US-09-198-452A-449 Sequence 449, App  
39 98 6.3 424 4 US-09-543-681A-4762 Sequence 4762, Ap  
40 97.5 6.2 336 4 US-09-489-039A-8430 Sequence 8430, Ap  
41 97.5 6.2 487 4 US-09-489-039A-8130 Sequence 8130, Ap  
42 97.5 6.2 551 4 US-09-252-991A-20358 Sequence 20358, A  
43 97 6.2 416 4 US-09-328-352-6537 Sequence 6537, Ap  
44 96.5 6.2 446 4 US-09-489-039A-9587 Sequence 9587, Ap  
45 96.5 6.2 467 4 US-09-107-532A-6175 Sequence 6175, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-252-991A-21214  
; Sequence 21214, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21214  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21214

Query Match 19.1%; Score 298; DB 4; Length 500;  
Best Local Similarity 33.3%; Pred. No. 9.9e-23;  
Matches 98; Conservative 38; Mismatches 94; Indels 64; Gaps 15;  
Qy 22 RDLKAAIAVGIGGVLLV-----GIVLS--PWGMYLVAGF-----MAAATWE 63  
Db 234 RIITALLVLPALGGFLLLEGAPFALFIGAVVSLGAWEARL-AGYEQQFGRVAYATVA 292  
Qy 64 VGRLEKGGVHLLPLMIGGQAIWLSWPFQGTWGLASFVATLVLMYFRIFYNGTEKE 123  
Db 293 V-----LMVALVHLSP---QLAGAVLLALVW-----WTALFVLVLY-----PE 328  
Qy 124 ARNYL--RDTSGVGFVLTWIPFGSPAAMLSLNNNSIPQTFILTFMLCVIASDYGGYI 181  
Db 329 SVGYWGRWRRLCGMLLILLPAWQGLVLLKQWPLANGL-----IIAVVLVWGADIGAYF 383  
Qy 182 AGVFFGSHPNAPVLSPKKSWEGPAGSIVLG-SVTGALSHPFLDHHWNG-----VILGCA 236  
Db 384 SGKAFGRKRLAPVSPFGKSWGYGGLAASLAITLAVGLY----RGWSGLAILLALGAA 439  
Qy 237 LVVCATL-GDLVESQFKRDLGIKDMNLLPGHGGLMDRLDGLMPA----AMVTW 285  
Db 440 LVVFSIVGDLTESMFKRSQGIKDSNLLPGHGGLVLDRLDLSLTAAIPVFAALLM 493

##### RESULT 2

US-09-489-039A-12390  
; Sequence 12390, Application US/09489039A

Patent No. 6610836  
 GENERAL INFORMATION:  
 APPLICANT: Gary Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709.2004001  
 CURRENT APPLICATION NUMBER: US/09/489,039A  
 PRIOR FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: US 60/117,747  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 14342  
 SEQ ID NO 12390  
 LENGTH: 289  
 TYPE: PRP  
 ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-12390

Query Match 18.9%; Score 295; DB 4; Length 289;  
 Best Local Similarity 29.3%; Pred. No. 9.6e-23;  
 Matches 85; Conservative 48; Mismatches 95; Indels 62; Gaps 11;  
 30 VGIIGLVLLGVLSPMG-----WYILVAGFMAATWEGSRKKEGGYHLPL 77  
 30 MGFAIVTVVCMALAEWQSGSTSTQRVWLAVLCGLLAAMLFL---LPEYHYDVHQ 86  
 78 PMITGGQAIWLSWPPFGTMGLASFVATVLMYFRFYNGTEKARNYLRDTSV--GI 135  
 87 PMV-----EGSLWAFAMW-----IVALLLVLSY-----PASAFAWRHRSKVLRLI 126  
 136 F-VLTWIPLF-GSFAAMLSLMMNSIPGTYFTILFMLCVIASDVGGYIAGVFFGSHPMAP 193  
 127 FGILITVFFWGLALRAHWADNHYSGALWLLYVWILVWGDADSGAYMFGKFKHKLAP 186  
 194 LVSPKSWEGFAGSIVLGSVGTGALSVEHFLDHHWMMG-----VILGALV--VCA 241  
 187 KVSFGKTVQGGFGLLTAAVIS-----WAYGVNHLVDTPTVLLVCSVVAALAS 235  
 242 TLGDLVESOFKEDIGIKMSNLLPGHGLMDRLDGMCPAAMVTWILSVI 291  
 236 VLGDJTESMFKREAGIKDGHILPGHGLDRIDSLTAAPVFACLLLIV 285

RESULT 3  
 US-09-282-218A-17  
 Sequence 17, Application US/09282218A  
 Patent No. 6503700  
 GENERAL INFORMATION:  
 APPLICANT: LEUNG, David W.  
 TITLE OF INVENTION: MAMMALIAN CDP-DIACYLGLYCEROL SYNTHASE  
 FILE REFERENCE: 077319/0153  
 CURRENT APPLICATION NUMBER: US/09/282,218A  
 CURRENT FILING DATE: 1999-03-31  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: Patent in ver. 2.0  
 SEQ ID NO 17  
 LENGTH: 249  
 TYPE: PRP  
 ORGANISM: E. coli  
 US-09-282-218A-17

Query Match 18.2%; Score 285; DB 4; Length 249;  
 Best Local Similarity 31.0%; Pred. No. 8.7e-22;  
 Matches 85; Conservative 38; Mismatches 87; Indels 64; Gaps 12;  
 47 WG-----WYILVAGFMAATWEGSRKKEGGYHLPLPIMITGGQAIWLS--- 91  
 7 WQQLSGFTTRSORVWLAVLCGLLALMLFL---LPEYHNTHQELVELS-----LWASLW 59  
 92 WFGTMGLASFVATVLMYFRFYNGTEKARNYLRDTSVGTIF-VLTWIPLF-GSFAA 149  
 60 W-----IVALLLVL-----FYPGSAATWRN--SKTLRLIFGVLTIVPFFWGLAL 102  
 150 MSLMNNNSIPGTYFTILFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIV 209

Db 103 RAWHYDENHYSYGAIMLLYVWILVWGDADSGAYMFGKLFGRHKLAPKVPCKTQWQFIGGLA 162  
 QY 210 LGSVTCALSVHFLDHHWMMGV-----ILGALV--VCATLGLDLSVESQFKRDLCI 257  
 Db 163 TAAVIS-----WYGMWANDVAPVILLCISVIAALASVLDLSTESMFKREAGI 211  
 QY 258 KMSNLLPGHGLMDRLDGMCPAAMVTWILSVI 291  
 Db 212 KDSCHLIPGHGILDRIDSLTAAPVFACLLLIV 245

RESULT 4  
 US-09-543-681A-6495  
 Sequence 6495, Application US/09543681A  
 Patent No. 6605709  
 GENERAL INFORMATION:  
 APPLICANT: GARY BRETON  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709.1002-001  
 CURRENT APPLICATION NUMBER: US/09/543,681A  
 CURRENT FILING DATE: 2000-04-05  
 PRIOR APPLICATION NUMBER: US 60/128,706  
 PRIOR FILING DATE: 1999-04-09  
 NUMBER OF SEQ ID NOS: 8344  
 SEQ ID NO 6495  
 LENGTH: 292  
 TYPE: PRP  
 ORGANISM: Proteus mirabilis  
 US-09-543-681A-6495

Query Match 18.1%; Score 282.5; DB 4; Length 292;  
 Best Local Similarity 29.0%; Pred. No. 2e-21;  
 Matches 88; Conservative 46; Mismatches 118; Indels 51; Gaps 10;  
 22 RDLKAAIAYGIGLVLLGIVLSPMGWYILVAGFMAATWE-----VGSRLKEG- 71  
 10 RVITALILIPVIALFL---LPPAGFGLVIAISGLAGWMAQFIGHWSQKRIATGV 65  
 72 GYHLPLPIM-----ITGGQAIWLSWPPFGTMGLASFVATVLMYFRIFY 117  
 66 GFAALVLMQASLPDPDHLDTAMIKNSLWAGLFW-----WVAAILLVWSYP--- 112  
 118 NGTEKEARNVLRDTSVGI FVLTWIPLF-GSFAAMLSLMMNSIPGTYFTILFMLCVIASD 176  
 113 NSANMKNVLLKLLPGI--LTIIVFYCGMMLRLMGLYSTNSYTGAWWLLYVWILVWAD 170  
 177 VGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIVLGSVGTGALSVEHFL---LDHWWMMGI 232  
 171 SGAYAFGLRMGXHKMAPKVPCKTLEGLVGLITAGVWSVWLFTRFAPITEVENHL---LL 227  
 233 LGCALVVCATLGLDLSVESQFKRDIGIKMSNLLPGHGLMDRLDGMCPAAMVTWILSVIS 292  
 228 ISGIVVIVSVFGDLAESMPKRSVKINDSSQLIPGHGVLDRIDSLTAAPVFACLLVWS 287  
 293 SSY 295  
 288 SGF 290

RESULT 5  
 US-09-328-352-7257  
 Sequence 7257, Application US/09328352  
 Patent No. 6562958  
 GENERAL INFORMATION:  
 APPLICANT: Gary L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 CURRENT FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252



SEQ ID NO 7257  
LENGTH: 285  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-7257

Query Match 16.8%; Score 262; DB 4; Length 285;  
Best Local Similarity 27.8%; Pred. No. 2.7e-19;  
Matches 82; Conservative 51; Mismatches 106; Indels 56; Gaps 8;

QY 19 NAGRLKAAIAVGIGLVLLGIVLS-----PWGWILVAGFMAAATW-----E 63  
DB 2 NSARENRIIMLERIVTALVAVVLGCMFATQSHYFMLVLMIVVAGVAGYKLMPRE 61  
QY 64 VQSRLEKGGYHLPDPMIIGQAIW-----LSWFFGTMGILASFVATVVLVY----- 112  
DB 62 VGAVVKPAWGYGLVAFVSGVALFFHDIALLLW-----SASILTWLVSVYWKSF 113  
QY 113 -PRIFYNGTEKARNYLRDTSVGIFVLTWPLFGSFAWLSLMQNNISPGTYFILTFLC 171  
DB 114 EFDGYNAT-----LVIGLILICAAVTAIFVWQSS-----FWMLMYLFL 155  
QY 172 VIADVGYYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHW---W 228  
DB 156 VWGADSGAYFVGRKFKLAPTSPKSVGLYGGILTTIIVMLVVOYQYVNLNLTWVQQL 215  
QY 229 MGVILGCALVCAITGLDIVEQFKRDIGIKDMSNLLPGHGGMDRLDQMLPAAMV 283  
DB 216 LFLILSLTVFSGVLGDLFESMIRKAGIKDSGRVLPGHGGVLDRIIDSLAAAPI 270

## RESULT 6

US-09-543-681A-7121  
Sequence 7121, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543.681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7121  
LENGTH: 315  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-7121

Query Match 16.3%; Score 255; DB 4; Length 315;  
Best Local Similarity 24.7%; Pred. No. 1.7e-18;  
Matches 73; Conservative 58; Mismatches 87; Indels 78; Gaps 13;

QY 26 AAIAGVIG-LGVLLGLIV-----LSP-----WGVIYLVAGFMAAATWEV 64  
DB 60 SAVAISIGTIGSVILFALISWALRELVTLPTRHGDHEALFWCFVV----- 107  
QY 65 GSRLKEGGYHLPDPMIIGQAIWLSWFFGTMGILASFVAT-VLVLMYPRIFYNGTEKE 123  
DB 108 -----LPIQYLVGVG---WYN-----LLAIFIPVYAFLIPIRMAISG---D 144  
QY 124 ARNYLR-----DTSVGIFVLTWPLFGSPAAMLSL-----MQNNISPGTYFILTFLC 172  
DB 145 THFLERMAKQVQSWMIATYCLVAP-----ALLMLPIEGFEGENI---NLLFLMIV 195  
QY 173 IASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHWNGVI 232  
DB 196 QTSVLQYVFGKLFGRKHPVPLKSPNKTIIEFFGGIISASLIG-MMLWATPTFWAAFL 254  
QY 233 LGCALVVCATIGDIVEQFKRDIGIKDMSNLLPGHGGMDRLDQMLPAAMVTLIL 288

DB 255 LSLVITILAGFAGGLCMASIKRDSGVKDFGTMIIEHGGMMDRMDSLCFAAPIFFHVI 310

## RESULT 7

US-09-540-236-2118  
Sequence 2118, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRI  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2118  
LENGTH: 296  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-2118

Query Match 15.9%; Score 248; DB 4; Length 296;  
Best Local Similarity 35.1%; Pred. No. 8.2e-18;  
Matches 74; Conservative 28; Mismatches 79; Indels 30; Gaps 8;

QY 89 WLSWPFGTWGIILASFVATVVLVLMYPRIFYNGTEKARNYLRDTSVGIFVLTWPLFGSFA 148  
DB 88 WVEFVAWSLGIWV--NATAWVAQY-----PNKERWYGRRLVYMGAVILT-----AAIT 133  
QY 149 AMLSLSMNNNSIPGTIFYILTFLCIVIASDVGVIAGVFFGSHPMAPLVSPKKSWEFGAGSI 208  
DB 134 AMYGLWQMS-----PWMLVYVFLVWCADSGAYFIGRKFGRKMAPNVSPKKSIEGLCGGL 189  
QY 209 -VLGSVTGALSVEHFLDHHWNGVILGCAL-----VVCATLGDIVESQFKRDIGIKDMSNL 263  
DB 190 MVSQVTVVAVGYLQLS---GMSLVFLILSLMTLVASILDGLFESMLKRRAGIKDSGTI 246  
QY 264 LPHGGGLMDRLDQMLPA-----AMVTVLILSV 290  
DB 247 LPHGGGLDRIDSLLSAIPFALGFWLLKL 277

## RESULT 8

US-09-134-001C-4966  
Sequence 4966, Application US/09134001C  
Patent No. 6380376  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4966  
LENGTH: 261  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4966

Query Match 15.8%; Score 247.5; DB 4; Length 261;  
Best Local Similarity 28.1%; Pred. No. 7.8e-18;  
Matches 79; Conservative 47; Mismatches 94; Indels 61; Gaps 10;

QY 22 RLKAAIAGVIGLVILGIVLSPGWVILVAGFMAAATWEVGSRLKEGGYHLPITMI 81  
DB 5 RTLTATIALILFPLIKGGLIL-----MLFALLALI-----ALKE-----L 42  
QY 82 IGGQAILWLSWPFGTWGIILASFVATVVLVLMYPRIFYNGTEKARNYLRDTSV-GIFVLTW 140

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5752
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5752

Query Match 14.0%; Score 219; DB 4; Length 192;
Best Local Similarity 40.0%; Pred. No. 4.9e-15;
Matches 48; Conservative 20; Mismatches 46; Indels 6; Gaps 2;

QY 165 ILTEMLCVI-ASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVMH-- 221
DB 61 VLLFALFIWADIDGAYLFGRRFGHKLMPDVPNKTIEGALGGILSAVVVAALFLVTA 120
QY 222 ---LLDHHWVGVLGCALVVCATLGDVLESQFKRDGKDKMSNLLPGHGLMDRLDGM 278
DB 121 NKGLFPYPMVPMVLTVLFSIVGQFDLVESIKRHYGVKDSGNILPGHGLDRFDSLL 180

RESULT 11
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Weeks, Reitha
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT
; SOFTWARE: WORD 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,814D
; FILING DATE: 28-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Drosophila
; STRAIN:
; INDIVIDUAL ISOLATE:

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31107
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31107

Query Match 15.8%; Score 246.5; DB 4; Length 346;
Best Local Similarity 25.3%; Pred. No. 1.5e-17;
Matches 80; Conservative 56; Mismatches 89; Indels 91; Gaps 16;

QY 33 GLGLVLL---GIVL-----SPMGWYILVAGFMAAATWVGS--- 66
DB 46 GIGALLASLVGVFLKWRAGPGPNVINDLNARINAWMVLTGP---AFW-LGQAAV 101
QY 67 -----RLKE-----GGYHLPBIMIGGQAIWLSWPFQWGL 100
DB 102 ILLFYGVSYALREFITLPTRRSDYPALVAAYFVLPQYL-----LIYDW-----YSMF 153
QY 101 ASFVAT-VIIVLMYFRIFYNGTEKEARNYLRDTS-----VGIFVLTWIPLFGSFAAMLS 152
DB 154 SIFIPYVFLLLPILASIGG---DTKHFLERASKVQWGLMAVFCISFVP-----ALLT 204
QY 153 LMQNSNIPG---TVFILTMLCVI-ASDVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGS 207
DB 205 L-----DIPGEGNLLIAYIVVQVSDVMQYICGKLFGRKIAPNLSPSKTVGFGV 260
QY 208 IVLGSVTGALSVMHFLDHHWVGVLGCALVVCATLGDVLESQFKRDGKDKMSNLLPGH 267
DB 261 IALATAIGA-SLWNTTPNPLWQALIALIINLLFGGLVMSAIKRGVGDWGHMIEGH 319
QY 268 GGLMDRLDGLMPLAAMV 283
DB 320 GGMLDRDLSVCFAPAI 335

RESULT 10
US-09-134-000C-5752
; Sequence 5752, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
```

```

;
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL LINE:
; CELL LINE:
; ORGANELLE:
;
; FEATURE:
; NAME/KEY: CDP-diacylglycerol synthase (CDS)
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
; US-08-672-814D-11
;
; Query Match 12.7%; Score 199; DB 2; Length 446;
; Best Local Similarity 23.3%; Pred. No. 1.9e-12;
; Matches 80; Conservative 40; Mismatches 111; Indels 112; Gaps 13;
;
; QY 50 YILVAGFAAATWEVGRSLKEGGYHLPPLIMI-----IGGQ-----AIIW 89
; DB 91 WIMICGF-----ALIIYGG---PLALMTLLVQVKCFQELIISIGYQVYRIHGLPW 138
;
; QY 90 ---LSWPFGTMGILAS--FVATVLVLMYFRIFYNGTE--KEARNYLRTDSVGIFVL--TW 140
; DB 139 FRSLSWYF---LLTSNYFFYGENLVDFYGVINRVEYLKFLVYTHRFLSFALYIIGFVW 194
;
; QY 141 IPL-----FGSPA-----AMLSIMQNNISPGTYFILTFLMCLVIASDVGGY 180
; DB 195 FVLSLVKKYIKQSLFANWTHSVLLIVTQSYLIIONIFEGLIWFIWVPSMIVCNDWY 254
;
; QY 181 IAGVFFGSHPMAPLVSPKKSWEAGFAGSIVLGSVTGALS VHFLLDHW----- 227
; DB 255 VFGFFGRTPLIKL-SPKKTWEGFIGGFAFVLFGLFSLVLCNYQYFICPIQYSEQR 313
;
; QY 228 -----WGVILGCVLVVCAITGDLVESQFKRD 254
; DB 314 MTMSCVPSYLFPTQEYSKLKLFGLGKTLNLYPFIWHSISLSLFSIIIGPFGFFASGFKRA 373
;
; QY 255 LGIKDMSNLLPGHGLMDRLDMLPAAMVTWLIISVSISSYPS 297
; DB 374 FKIKDFGDMIPGHGIMDRPDCQFLMATFVNVIISFIRTPSPA 416
;
; RESULT 12
; US-09-333-696-11
; Sequence 11, Application US/09333696
; Patent No. 6200769
;
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Weeks, Reitha
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS version 6.1, Windows NT
; SOFTWARE: WORD 6.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,696
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/672,814
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Drosophila
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
;
; FEATURE:
; NAME/KEY: CDP-diacylglycerol synthase (CDS)
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
; US-09-333-696-11
;
; Query Match 12.7%; Score 199; DB 3; Length 446;
; Best Local Similarity 23.3%; Pred. No. 1.9e-12;
; Matches 80; Conservative 40; Mismatches 111; Indels 112; Gaps 13;
;
; QY 50 YILVAGFAAATWEVGRSLKEGGYHLPPLIMI-----IGGQ-----AIIW 89
; DB 91 WIMICGF-----ALIIYGG---PLALMTLLVQVKCFQELIISIGYQVYRIHGLPW 138
;
; QY 90 ---LSWPFGTMGILAS--FVATVLVLMYFRIFYNGTE--KEARNYLRTDSVGIFVL--TW 140
; DB 139 FRSLSWYF---LLTSNYFFYGENLVDFYGVINRVEYLKFLVYTHRFLSFALYIIGFVW 194
;
; QY 141 IPL-----FGSPA-----AMLSIMQNNISPGTYFILTFLMCLVIASDVGGY 180
; DB 195 FVLSLVKKYIKQSLFANWTHSVLLIVTQSYLIIONIFEGLIWFIWVPSMIVCNDWY 254
;
; QY 181 IAGVFFGSHPMAPLVSPKKSWEAGFAGSIVLGSVTGALS VHFLLDHW----- 227
; DB 255 VFGFFGRTPLIKL-SPKKTWEGFIGGFAFVLFGLFSLVLCNYQYFICPIQYSEQR 313
;
; QY 228 -----WGVILGCVLVVCAITGDLVESQFKRD 254
; DB 314 MTMSCVPSYLFPTQEYSKLKLFGLGKTLNLYPFIWHSISLSLFSIIIGPFGFFASGFKRA 373
;
; QY 255 LGIKDMSNLLPGHGLMDRLDMLPAAMVTWLIISVSISSYPS 297
; DB 374 FKIKDFGDMIPGHGIMDRPDCQFLMATFVNVIISFIRTPSPA 416
;
; RESULT 13
; US-09-282-218A-19
; Sequence 19, Application US/09282218A
; Patent No. 6503700
;
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; TITLE OF INVENTION: MAMMALIAN CDP-DIACYLGLYCEROL SYNTHASE
; FILE REFERENCE: 077319/0153
; CURRENT APPLICATION NUMBER: US/09/282,218A
; CURRENT FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 446

```

```
; TYPE: PRT
; ORGANISM: Drosophila
US-09-282-218A-19
Query Match      12.7%; Score 199; DB 4; Length 446;
Best Local Similarity 23.3%; Pred. No. 1.9e-12;
Matches 80; Conservative 40; Mismatches 111; Indels 112; Gaps 13;

QY 50 YILVAGFMAAATWVGSRRLKEGGYHLPIMT-----IGGG-----AIIW 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 WIMICGF-----ALIIYCG-----PLALMITLLVQVKCFQBIISIGYQYRIHGLPW 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 ----LSWPFCTMGILAS--FVATVLVMTFRFYNGTE--KEARNVLRDTSVGIIVL--TW 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 FRSLSWYF-----LLTSNFFYGENLDVYGVVNRVEYKFLVTHRFSLPALYIIFGW 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 141 IPL-----FGSEA-----AMLSLMQNNISIPGTFFILTMCLCVIASDVGGY 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 FVLSLVKKYIKQFSLPFWTHVSLIIVTQSVLIIONIFESGLIWFIVPVSIVCNDWVAY 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 IAGVFEGSHMAPLVSPKKSWEFGFAGSIVLGSVTGALSVMHFLDHHW-----227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 VGFPEFGRTPLIKL-SPKKTWEGFIGGGFATVLPGLFYSVYLCNVQYFICPIQXSEBQR 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 -----NMGVILGCALVVCATLGLDVLVESQEKD 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 314 WTMSCVPYSYLPQBYSKLFGIGKTLNLYPFPIHSHISLSLFSLSIIGPFGGFASGFRA 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 LGIKDMSNLLPGHGLMDRLDGLMPLAAMVTLIISVSISSYPS 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 374 FKIKDFGDMIPGHGIMDRFDCQFLMATFVNVIYSFIRTPSPA 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-282-218A-12
; Sequence 12, Application US/09282218A
; Patent No. 6503700
; GENERAL INFORMATION:
; APPLICANT: LEUNG, David W.
; TITLE OF INVENTION: MAMMALIAN CDP-DIACYLGLYCEROL SYNTHASE
; FILE REFERENCE: 077319/0153
; CURRENT APPLICATION NUMBER: US/09/282.218A
; CURRENT FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-282-218A-12
Query Match      12.7%; Score 198.5; DB 4; Length 445;
Best Local Similarity 27.5%; Pred. No. 2.2e-12; Indels 125; Gaps 19;
Matches 98; Conservative 32; Mismatches 101; Indels 141; Gaps 20;

QY 32 IGLGVLLGLVILGVLSPGWYIILVAGFMAAATWVGSRRLKEGGYHLPIMTIGGQAIW-- 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 IYLGFWLMIIVMC-----VQIKCFHEIIT--IGYNVVH-SYDLP-----WFR 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 -LSWPFCTMGILASFVATVLVMTFRFYNGT-----EKEAR---NYLRDTS-- 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 TLSWYF-----LLCVNYF--FYGTVTVDYFTLVQREEPRLISKYHRFISFT 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 ---VG--IFVLMTIP-----LFG-SFRAML-----SLMQNNSIPGTFFILTMCLV 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 LYLIGFCFVLSLVKKYHVLQYVFGFWGTHVTLIIVVQSHLVHNLFEQMIWFTVPIVSCV 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 IASDVGGYIAGVFGSHPMAPLVSPKKSWEFGFAGSIVLGSVTGAL-----217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 233 ICNDIMAYMFGFFFGRTPLIKL-SPKKTWEGFIGGFATVWFGLLSYVMGSRVCFVCPV 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 218 -----SVHFLLDHWHMMGVIL-----GCALVVCATL---243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

292 EYNNDTNSFTVDCPSDLFRLOEYNIQVIGWKTVMYPPFQIHSTALSTFASLIGP 351
244 -GDLVESQPKRDLGIDKMSNLLPGHGLMDRLD-GMLPAAVMTWLILSVISSYPS 297
352 FGGFFASGFKRAFKINDPANTIPGHGIMDRFDCQYLMTATFVNVIASIFRGNFS 407

RESULT 15
US-08-672-814D-2
; Sequence 2, Application US/08672814D
; Patent No. 5952480
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Weeks, Reitha
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT
; SOFTWARE: WORD 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,814D
; FILING DATE: 28-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; HYPOTHETICAL: no
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: CDS
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-672-814D-2
Query Match      12.7%; Score 198.5; DB 2; Length 461;
Best Local Similarity 24.8%; Pred. No. 2.3e-12; Mismatches 113; Indels 141; Gaps 20;
Matches 98; Conservative 43; Mismatches 113; Indels 141; Gaps 20;

QY 13 MPKPKNAGRDLEKAAIYVIGL-----GVLVL-----LGIVLSPGWYIILVAG 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 IPSSDRTPKILKALS---GLSRKWKWIRGILTMISLFFLIIMGSMMLMLVLG 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 56 FMAAATWE---VGSRLKEGGYHLPLPIMIIGGOAILM---LSWPFGTMGILASEVATVLV 109  
Db 118 IOVKCFHEIITIGRYVH-SYDLP-----WFRLESWYF-----LLC 152  
QY 110 LMYFRIFYNGT-----EKEAR-----NYLRDTSVGI-----FVLTWIP-----L 143  
Db 153 VNYF--PYGETVADYFATFVOREEQQLIRYHRFISFALYLAGPCMFVLSLVKEHYRLQ 210  
QY 144 FGSFA-----AMLSLQN-----NSIPGTYFILTFLMLCVIASDVGGYIAGVREGSHPMAP 193  
Db 211 FYMFANWHTVLLITVTQSHLVIQNLFEGMWELVLPVISSVICNDITAYLFGFFFGRTPLIK 270  
QY 194 LVSPKKSWEFGFAGSIVLGSVTGALSVMFLDHHWM-----GVILGC----- 235  
Db 271 L-SPKKTWEGFIGGFFTVVFGFIAAVLSKYQYFVCPVEYRSDVNSFVTECPSELFOL 329  
QY 236 -----ALVVCATL-----GDLVESQFKRDIGIKDMSNL 263  
Db 330 QTYSLPPLXAVLRQERSVLYPQIHSLALSTFASLIGPFGPFASGFKRAFKIKDFANT 389  
QY 264 LPHGGGIMDRLD-CMLPAAWTVLILSVISSYPS 297  
Db 390 IPHGGIMDRFDCOYLMATLVHGYIITSFIRGNPS 424

Search completed: August 11, 2004, 13:36:57  
Job time : 20 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:46:53 ; Search time 8196 Seconds  
(without alignments)  
6874.810 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagtgctgtcgcaagga.....aggaaccgattccggctgg 1300

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1300	100.0	1300	6	AX297999	Sequence
C 2	1300	100.0	337200	1	AP005280	Coryneb
C 3	1300	100.0	349980	6	AX127149	Sequence
C 4	1300	100.0	349980	6	AX127150	Sequence
C 5	1193	91.8	1194	1	CGJ312009	Coryneb
C 6	978	75.2	978	6	AX064835	Sequence
C 7	978	75.2	978	6	AX066517	Sequence
C 8	855	65.8	855	6	AX122300	Sequence
C 9	855	65.8	855	6	BD164417	Novel pol
C 10	596.8	45.9	303250	1	AP005220	Coryneb
C 11	342.2	26.3	348408	1	BX248358	Coryneb
C 12	191.8	14.8	16952	1	AE007118	Corybacte
C 13	191.8	14.8	318050	1	BX248344	Mycobacte
C 14	191.8	14.8	348676	15	BX842581	Myco
C 15	164	12.6	36985	1	MSGB1529CS	Myco
C 16	164	12.6	318200	1	MLPRTN6	Myco
C 17	123.4	9.5	297	6	AX307920	Sequence
C 18	117	9.0	258	6	AX122299	Sequence
C 19	117	9.0	258	6	BD164416	Novel pol
C 20	110.2	8.5	299550	1	AP005031	Streptomy
C 21	107	8.2	308050	1	SC0939124	Streptomy
C 22	75.6	5.8	349981	1	BX572602	Rhodopseu
C 23	73.8	5.7	10029	1	AE012236	Xanthomon
C 24	72.4	5.6	300880	1	AE016917	Chromobac
C 25	72	5.5	299050	1	AE016917	Tropherym
C 26	72	5.5	302529	1	AE016851	Tropherym
C 27	71.8	5.5	10242	1	AE011773	Xanthomon
C 28	69.8	5.4	10771	1	AE014782	Bifidobac
C 29	69.8	5.4	349980	6	AX492787	Sequence
C 30	69.8	5.4	349980	6	AX553954	Sequence
C 31	69.6	5.4	300600	1	AP005352	Bradyrhiz
C 32	68.4	5.3	310029	1	AE016861	Pseudomon
C 33	67.2	5.2	301995	1	AE016779	Pseudomon
C 34	65.8	5.1	1845	1	PSECDEDS	Pseudomonas
C 35	65.8	5.1	14537	1	AE004785	Pseudomon
C 36	65	5.0	11455	1	AE012802	Chlorobiu
C 37	64.4	5.0	295250	1	BX321862	Nitrosomo
C 38	64	4.9	342805	1	BX571861	Photorhab
C 39	64	4.9	349980	6	AX770908	Sequence
C 40	63.8	4.9	300817	1	AE016943	Bacteroid
C 41	62.8	4.8	299350	1	AP006572	Gloeobact
C 42	60.8	4.7	10487	1	AE003942	Xylella f
C 43	60.8	4.7	300029	1	AE012554	Xylella f
C 44	59	4.5	10301	1	AE004297	Vibri
C 45	58.8	4.5	11002	1	AE004681	Pseudomon

ALIGNMENTS

RESULT 1	AX297999	AX297999	1300 bp	DNA	linear	PAT 26-NOV-2001
LOCUS	Sequence 1 from Patent WO0193785.					
DEFINITION	AX297999					
ACCESSION	AX297999					
VERSION	AX297999.1	GI:17128085				
KEYWORDS	Corynebacterium glutamicum					
SOURCE	Corynebacterium glutamicum					
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.					
REFERENCE	1					
AUTHORS	Nampoothiri, M., Moeckel, B., Pfeifferle, W., Eggeling, L. and Sahm, H.					
TITLE	Nucleotide sequences coding for the cda gene					
JOURNAL	Patent: WO 0193785-A 1 98-NOV-2001					

FEATURES		Degussa AG (DB) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)	
source	1..1300	Location/Qualifiers	
		/organism="Corynebacterium glutamicum"	
RBS	187..195	/db_xref="taxon:1718"	
		/gene="cdsA"	
gene	200..1093	/gene="cdsA"	
		/codon_start=1	
CDS	200..1093	/transl_table=11	
		/protein_id="CAD12679.1"	
ORIGIN	100.0%; Score 1300; DB 6; Length 1300;	Query Match	
	Best Local Similarity 100.0%; Pred. No. 0;	Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GAAATCGTGTGTCGCAAGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 60	
DB	1	GAAATCGTGTGTCGCAAGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 60	
QY	61	GCTTAAGAAATACGTCGTTCAAGTCGATGGCGGGGAAACACGCTTCTTTTCGCTT 120	
DB	61	GCTTAAGAAATACGTCGTTCAAGTCGATGGCGGGGAAACACGCTTCTTTTCGCTT 120	
QY	121	GCAAGAGTGTGTTGGAAGAAATTTTTCGAAAATGCTGGCACCACATCAACAGTGCATTTGTA 180	
DB	121	GCAAGAGTGTGTTGGAAGAAATTTTTCGAAAATGCTGGCACCACATCAACAGTGCATTTGTA 180	
QY	181	GAACTTCAAGAGAACCCATGAATGAACCGAGCAACATCACCGTTCATGAGATGCC 240	
DB	181	GAACTTCAAGAGAACCCATGAATGAACCGAGCAACATCACCGTTCATGAGATGCC 240	
QY	241	CAAAACCAAAATATGCGGTGCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACT 300	
DB	241	CANACCAAAATATGCGGTGCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACT 300	
QY	301	GGGGTCTCTGTTCTTTTGGGATGTCCTAAGCCCATGGGGTGGTACATCTCTGTTGC 360	
DB	301	GGGGTCTCTGTTCTTTTGGGATGTCCTAAGCCCATGGGGTGGTACATCTCTGTTGC 360	
QY	361	AGTTTATGCTGCAGCAACATCGGAGTTGGTAGAGACTTAAAGAGCGGCTATCA 420	
DB	361	AGTTTATGCTGCAGCAACATCGGAGTTGGTAGAGACTTAAAGAGCGGCTATCA 420	
QY	421	TTTGCACTGCGGATATGATCATCGGCGGTGAGGCAATCATCTGGCTGATGATTCGG 480	
DB	421	TTTGCACTGCGGATATGATCATCGGCGGTGAGGCAATCATCTGGCTGATGATTCGG 480	
QY	481	TGGCAGATGGCAATTTTGGGCTCTTTTGGGCACTGTGTGGTGTGATGATTTCCG 540	
DB	481	TGGCAGATGGCAATTTTGGGCTCTTTTGGGCACTGTGTGGTGTGATGATTTCCG 540	
QY	541	AATTTTCTAATGCGACGAAAAGAGCCCGCAACTATTATGAGGACACCTCTGTGGG 600	
DB	541	AATTTTCTAATGCGACGAAAAGAGCCCGCAACTATTATGAGGACACCTCTGTGGG 600	
QY	601	CATCTTCTGCTCAGCTGATTCATTTTGGAGCTTCGCTCGGATGCTGCTGAT 660	
DB	601	CATCTTCTGCTCAGCTGATTCATTTTGGAGCTTCGCTCGGATGCTGCTGAT 660	
QY	661	GCAAAACAATCCATCCCGGGTACATATTTCATTTTGCAGTTCATGCTGTGTGATCGC 720	

Db	661	GCAAAACAATTCATCCGGGTACATATTTCATTTGACGTTTCATGCTGTGTGATCGC	720
Qy	721	ATCGGATGTGGCGGGTATATCGCGGTGTGTCTTTTGGATCGCACCAATCGCGCGTT	780
Db	721	ATCGGATGTGGCGGGTATATCGCGGTGTGTCTTTTGGATCGCACCAATCGCGCGTT	780
Qy	781	GGTAGTCCGAAGAAGTCTTTGGAAAGGCTTTCGGGCTCCATTGCTTAGGATCGGTAC	840
Db	781	GGTAGTCCGAAGAAGTCTTTGGAAAGGCTTTCGGGCTCCATTGCTTAGGATCGGTAC	840
Qy	841	TGGTGCACTCAGTGTTCATCTCTGCTCGATCACCATCGGTGGTGGTGTGATCTTGGG	900
Db	841	TGGTGCACTCAGTGTTCATCTCTGCTCGATCACCATCGGTGGTGGTGTGATCTTGGG	900
Qy	901	TTTGCCCTAGTGTGTGCGGCACCTTGGGTGACTTGGTTGAGTCGAGTTCGAAACGCGA	960
Db	901	TTTGCCCTAGTGTGTGCGGCACCTTGGGTGACTTGGTTGAGTCGAGTTCGAAACGCGA	960
Qy	961	TTTGGGCATCAAGGATATGTCGAACTTCTTTCAGGCGCACGGCGGATGATGACCGTTT	1020
Db	961	TTTGGGCATCAAGGATATGTCGAACTTCTTTCAGGCGCACGGCGGATGATGACCGTTT	1020
Qy	1021	GGATGCACTGCTCCCGCGCGGATGCTGAGTGGTTGATCTGAGTGTGATCAGCAGCTC	1080
Db	1021	GGATGCACTGCTCCCGCGCGGATGCTGAGTGGTTGATCTGAGTGTGATCAGCAGCTC	1080
Qy	1081	GTATCGCTGTAAGCTTGGCGCAGCTTTAAAGTTCAAAAACCTTGAAGGCGCTGAGGTG	1140
Db	1081	GTATCGCTGTAAGCTTGGCGCAGCTTTAAAGTTCAAAAACCTTGAAGGCGCTGAGGTG	1140
Qy	1141	CATAACGCTGCACTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGCTTTTACGGATTT	1200
Db	1141	CATAACGCTGCACTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGCTTTTACGGATTT	1200
Qy	1201	TTCTTAACGTCGACGCTACTCAAAATGCGGACCGCCCAACAGGCGCCATATCAATG	1260
Db	1201	TTCTTAACGTCGACGCTACTCAAAATGCGGACCGCCCAACAGGCGCCATATCAATG	1260
Qy	1261	CACCGTAAATGCTGCTAGTAGGAACCGGATTCGGGCTGG	1300
Db	1261	CACCGTAAATGCTGCTAGTAGGAACCGGATTCGGGCTGG	1300
RESULT 2			
AP005280/c			
LOCUS	337200 bp	DNA	linear BCT 08-AUG-2002
DEFINITION	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 7/10.		
ACCESSION	AP005280	BA000036	
VERSION	AP005280.2	GI:22138850	
KEYWORDS	Corynebacterium glutamicum ATCC 13032		
SOURCE	Corynebacterium glutamicum ATCC 13032		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1		
AUTHORS	Nakagawa, S.		
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 337200)		
AUTHORS	Nakagawa, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@anagen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)		
COMMENT	On Aug 8, 2002 this sequence version replaced gi:21324668. This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.		
FEATURES	Location/Qualifiers		
source	1..337200		
	/organism="Corynebacterium glutamicum ATCC 13032"		



/mol\_type="genomic DNA"  
 /strain="ATCC 13032"  
 /db\_xref="taxon:196627"  
 /note="ATCC 13032"  
 complement(518..1216)  
 /gene="Cg11899"  
 complement(518..1216)  
 /gene="Cg11899"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Uncharacterized ACR"  
 /protein\_id="BAB99292.1"  
 /db\_xref="GI:21324671"  
 translation="MSELDIKQLNKLQRYSQWAVRAIPGALDDDRRTVTVQAAKFFA  
 DLKACKTVIRGINASGRADADNIMWHAEEFDIOKAFADPRTILIQGVSEVFW  
 AVEFDFRANTVPEFALGDYEWLVAFADELHRIYDLMHKRYTEARLEVEELPFI  
 GORVDIADIKVLP"  
 1489..2115  
 /gene="Cg11900"  
 /gene="Cg11900"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Hypothetical protein"  
 /protein\_id="BAB99293.1"  
 /db\_xref="GI:21324670"  
 translation="MQIVDSFATSCHKTSATPAESTAESEAVESMRAELPELT  
 IGTIPRQRLAPFSAIGLEVQNGESDDVINSRGSFGRLILLHDPGAEETWEGAM  
 RLIVATQDMHVASDELLPEVAQWNGLEQAGATNLGGTVITSTTVRFGEIG  
 GPISAYQVEMASWTATGTLTAHVEAFVAALVASAGLPPGVTGLR"  
 2072..3337  
 /gene="Cg11901"  
 2072..3337  
 /gene="Cg11901"  
 /EC\_number="3.1.26.3"  
 /note="PF00570:HRDC domain  
 PF01612:3'-5' exonuclease"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Ribonuclease D"  
 /protein\_id="BAB99294.1"  
 /db\_xref="GI:21324671"  
 translation="WLDFFQASPNYGRGLTWMSDLLQPRDGPPIPLSTPGFTAA  
 DLLASGTGFADITERAGFYDRAFLIQRSGSTLFDPQFRPELTQALKPVL  
 NGQWIIHAASDLPFLAWLDHPGLLPDTLACRLAGFDHVNLAAMVEQFIDLHLK  
 GHSDDWKRPLPESNLVAALDVMLELADVAEILDDQOKLPWABOEPRFVVDQF  
 ATMTPESTMDQLKGLSTLKPDPOLVAREWMLERDSFASRLDAPCKVLSKVIYE  
 VARVLPRTPALQAQKGPFGSQSKATKWFPIITRALKSPRRNFKPQQRKDGIDRR  
 AWASTYPEBEHLQETRALIDDLAADINPGENILQPSLTVAVVMMAKHTGEIHAET  
 LNAVLDYGAQWQIDQFFPILSANLLKL"  
 complement(3399..5309)  
 /gene="Cg11902"  
 complement(3399..5309)  
 /gene="Cg11902"  
 /note="PF02779:Transketolase, pyridine binding domain  
 PF02780:Transketolase, C-terminal domain  
 TIGR00204:oxs: deoxyxylulose-5-phosphate synthase"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Deoxyxylulose-5-phosphate synthase"  
 /protein\_id="BAB99295.1"  
 /db\_xref="GI:21324672"  
 translation="MGLINLSITPADLKALNDEDLALAKEIRITFLVDKVAATCGHLG  
 PNLGVELTILGHVAFSDPQPIFDTSQHSQYVHKILTGRAXDFDSLRQDGLSGYTC  
 RAESEHDTVSHSAALSADGKSKALQDGDTHSVAVVGDGALTGGMCFEALNN  
 IAAGRKRVVVNDNGRSYPTTGGFAENLAGLMQFPYDFRMEKGTSLSKSMGVG  
 TETFEALHAFKEGVKSTVPTTMEPRLGMKVGVGDVGHNOKAVDNALKYADHDGPII  
 VHMTEKGRGYAPABQDLDELHMGSTGVDLTGAPKSAKPGWTSVSDELVKTAGON  
 ENVVAITAAAGPTGLSKFEANFNPRFDVGLAQCHAVTSAGLALGCKHEVVAIYT  
 FLNRAFOQLMDGMNLQFVLVLRDSRPSFGDGSNGHVMGMALTSVLPGVQAAPR  
 DEDSLRLNEAISIDGDTVVRFPKGLPFIIVAIPTLEDGVDLAYEDATDVST  
 DAPSVLITAVGERATVLDVASRIKQHGWNVTVDPRWIVPQSLVALSDHDLVIT

/product="Hypothetical protein"  
 /protein\_id="BAB99300.1"  
 /db\_xref="GI:21324677"  
 /translation="MNINHTYQVSWSEDDQEFATVAEPFSLSWLFSQDQAEABAGLR  
 NLVAEVLDMKQTGEQIPTSVGRNL"  
 complement (9277..9570)  
 /gene="Cg11908"  
 CDS  
 complement (9277..9570)  
 /gene="Cg11908"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Hypothetical protein"  
 /protein\_id="BAB99301.1"  
 /db\_xref="GI:21324678"  
 /translation="MADYDPRRRVDELETSLEGKAVENANSMDDGRIVBSF  
 BIPNVLSGEELNVVPRRADEFTCASCFILVQRNRKSHVEPDGSIICLDC"  
 complement (9721..10539)  
 /gene="Cg11909"  
 CDS  
 complement (9721..10539)  
 /gene="Cg11909"  
 /note="PF00459:Inositol monophosphatase family"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Archaeal fructose-1,6-bisphosphatase and related  
 enzymes of inositol monophosphatase family"  
 /protein\_id="BAB99302.1"  
 /db\_xref="GI:21324679"  
 Query Match 100.0%; Score 1300; DB 1; Length 337200;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAAGTCGTTGTCGCAAGAAAAGAACTGATGAGGTCCTAGAGACCTTTATCGCAATG 60  
 DB 137324 GAAGTCGTTGTCGCAAGAAAAGAACTGATGAGGTCCTAGAGACCTTTATCGCAATG 137265  
 QY 61 GCTTAAGAAATACGTCGTTCAACGTCGATGAGGCGGGGAAACGACGCTTCTTTGCTT 120  
 DB 137264 GCTTAAGAAATACGTCGTTCAACGTCGATGAGGCGGGGAAACGACGCTTCTTTGCTT 137205  
 QY 121 GCAAGAGTGTGTTGGAAGAAATTTTTCGAAATGCTGCGACCATCAACAGTGACATGTTGA 180  
 DB 137204 GCAAGAGTGTGTTGGAAGAAATTTTTCGAAATGCTGCGACCATCAACAGTGACATGTTGA 137145  
 QY 181 GAAATCTCAAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCATGAGATGCC 240  
 DB 137144 GAAATCTCAAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCATGAGATGCC 137085  
 QY 241 CAAACCCAAAATAATGCGGTGCGAGATCTCAAGCTGCCATTCCTGTGGGATCGGACT 300  
 DB 137084 CAAACCCAAAATAATGCGGTGCGAGATCTCAAGCTGCCATTCCTGTGGGATCGGACT 137025  
 QY 301 GCGGCTCCTGCTTCTTTGGGATGTCCTAAGCCCATGGGTTGGTACATCCTCGTTGC 360  
 DB 137024 GCGGCTCCTGCTTCTTTGGGATGTCCTAAGCCCATGGGTTGGTACATCCTCGTTGC 136965  
 QY 361 AGGTTTATGCTGAGCAACATGGAAGATGTTGTTAGCAGACTTAAGAGCGGCTATCA 420  
 DB 136964 AGGTTTATGCTGAGCAACATGGAAGATGTTGTTAGCAGACTTAAGAGCGGCTATCA 136905  
 QY 421 TTGTCACCTGCCGATTATGATCATCGCGGTGAGGCAATCATCTGGCTGTATGCCATT 480  
 DB 136904 TTGTCACCTGCCGATTATGATCATCGCGGTGAGGCAATCATCTGGCTGTATGCCATT 136845  
 QY 481 TGGCAGATGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTGTGATGATTTCCG 540  
 DB 136844 TGGCAGATGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTGTGATGATTTCCG 136785  
 QY 541 AATTTTCTACATGACCGGAAAAGACCGGCAACTATTGAGGACACCTCTGTGGG 600  
 DB 136784 AATTTTCTACATGACCGGAAAAGACCGGCAACTATTGAGGACACCTCTGTGGG 136725  
 QY 601 CATCTCGTGTACCTGGATTCCATTGTTGGAAAGCTTCGTCGATGCTGTCTGAT 660

DB 136724 CATCTCGTGTCACTGGATTCCATTGTTGGAAAGCTTCGTCGATGCTGTCTGAT 136665  
 QY 661 GCAAAACAAATTCATCCCGGTACATATTTTCAATTTTGAGCTTCATGCTGTGTGATCGC 720  
 DB 136664 GCAAAACAAATTCATCCCGGTACATATTTTCAATTTTGAGCTTCATGCTGTGTGATCGC 136605  
 QY 721 ATCGGATGTGGCGGGTATATCGCGGTGTGTCTTTTGATCGCACCAATGCGGCCGCTT 780  
 DB 136604 ATCGGATGTGGCGGGTATATCGCGGTGTGTCTTTTGATCGCACCAATGCGGCCGCTT 136545  
 QY 781 GGTAGTCCGAAGAAGTCTTTGGGAAGCTTTGCGCGCTCCATTTCTTAGATCGGTAC 840  
 DB 136544 GGTAGTCCGAAGAAGTCTTTGGGAAGCTTTGCGCGCTCCATTTCTTAGATCGGTAC 136485  
 QY 841 TGTGCACTCACTGTTTCACTTCTCTGATCAACCACTGTTGATGTTGATGTTGATCTTGG 900  
 DB 136484 TGTGCACTCACTGTTTCACTTCTCTGATCAACCACTGTTGATGTTGATGTTGATCTTGG 136425  
 QY 901 TTGTGCCCTAGTTGTGTGCGCCACGTTGGGTGACTTTGAGTCCAGTTCAAACGCGA 960  
 DB 136424 TTGTGCCCTAGTTGTGTGCGCCACGTTGGGTGACTTTGAGTCCAGTTCAAACGCGA 136365  
 QY 961 TTGGGCATCAAGGATATGTCGAACCTTCTCCAGCCACGCGGATTTGATGACCGTTT 1020  
 DB 136364 TTGGGCATCAAGGATATGTCGAACCTTCTCCAGCCACGCGGATTTGATGACCGTTT 136305  
 QY 1021 GGATGTCATGCTCCCGCGCGCATGTTGAGTGGTGTGATCTCTGATGATCAGCAGTCT 1080  
 DB 136304 GGATGTCATGCTCCCGCGCGCATGTTGAGTGGTGTGATCTCTGATGATCAGCAGTCT 136245  
 QY 1081 GTATCGCTGTTAAGCTTTGGGCGAGCTTTAAGTTCAAAAACCTTCAAAGGCGCTGAGTG 1140  
 DB 136244 GTATCGCTGTTAAGCTTTGGGCGAGCTTTAAGTTCAAAAACCTTCAAAGGCGCTGAGTG 136185  
 QY 1141 CATACGCTGTGCACTCAAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGGATTT 1200  
 DB 136184 CATACGCTGTGCACTCAAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGGATTT 136125  
 QY 1201 TTCTTAACCTGGCGAGGTTACTCAACATGCGCACGCCCAACAAAGCGCCATATCAATG 1260  
 DB 136124 TTCTTAACCTGGCGAGGTTACTCAACATGCGCACGCCCAACAAAGCGCCATATCAATG 136065  
 QY 1261 CACCGTATGCTGCTAGTAGGAACCGATTCCGCTGG 1300  
 DB 136064 CACCGTATGCTGCTAGTAGGAACCGATTCCGCTGG 136025  
 RESULT 3  
 AX127149/c 349980 bp DNA linear PAT 11-MAY-2001  
 LOCUS  
 DEFINITION Sequence 7065 from Patent EP1108790.  
 AX127149 AX114121  
 ACCESSION  
 VERSION AX127149.1 GI:14041137  
 KEYWORDS  
 SOURCE Corynebacterium glutamicum  
 ORGANISM  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 REFERENCE 1  
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.  
 TITLE Novel polynucleotides  
 JOURNAL Patent: EP 1108790-A 7065 20-JUN-2001;  
 KYOWA HAKKO KOGYO CO., LTD. (JP)  
 FEATURES  
 source Location/Qualifiers  
 1..349980  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1718"  
 /note="Seq 1 to long (3.309.400) split in 11, seq 7065  
 1.800.001 2.149.980"

ORIGIN

Query Match 100.0%; Score 1300; DB 6; Length 349980; Best Local Similarity 100.0%; Pred. No. 0; Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG	60						
DB	337324	GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG	337265						
QY	61	GCTTAAGAAATACGTGCTTTCAAGCTGATGGCGGGGAAAAGACGCTTTCTTTGCTT	120						
DB	337264	GCTTAAGAAATACGTGCTTTCAAGCTGATGGCGGGGAAAAGACGCTTTCTTTGCTT	337205						
QY	121	GCAAGAGTGTGGAGAAATTTTTTCAGAAATGCTGCACCATCAACAGTGAATGTTA	180						
DB	337204	GCAAGAGTGTGGAGAAATTTTTTCAGAAATGCTGCACCATCAACAGTGAATGTTA	337145						
QY	181	GAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGC	240						
DB	337144	GAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGC	336065						
QY	241	CAAAACCAAAATATCGGGTCTGAGATCTCAAGCTGCATTTGCTGTGGGATCGGACT	300						
DB	336064	CAAAACCAAAATATCGGGTCTGAGATCTCAAGCTGCATTTGCTGTGGGATCGGACT	336305						
RESULT 4									
AX127150/c 349980 bp DNA linear PAT 11-MAY-2001									
Sequence 7066 from Patent EP1108790.									
AX127150 AX114121									
AX127150.1 GI:14041138									
Corynebacterium glutamicum									
Corynebacterium glutamicum									
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;									
Corynebacterineae; Corynebacteriaceae; Corynebacterium.									
REFERENCE									
1									
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,									
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.									
TITLE Novel polynucleotides									
JOURNAL Patent: EP 1108790-A 7066 20-JUN-2001;									
KIOWA HAKKO KOGYO CO., LTD. (JP)									
FEATURES									
1. 349980									
Location/Qualifiers									
/organism="Corynebacterium glutamicum"									
/mol_type="genomic DNA"									
/db_xref="taxon:1718"									
/note="Seq 1 to long (3,309,400) split in 11, seq 7066									
2,100,001 2,449,980"									
ORIGIN									
Query Match 100.0%; Score 1300; DB 6; Length 349980;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG	60						
DB	337324	GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG	337265						
QY	61	GCTTAAGAAATACGTGCTTTCAAGCTGATGGCGGGGAAAAGACGCTTTCTTTGCTT	120						
DB	337264	GCTTAAGAAATACGTGCTTTCAAGCTGATGGCGGGGAAAAGACGCTTTCTTTGCTT	337205						
QY	121	GCAAGAGTGTGGAGAAATTTTTTCAGAAATGCTGCACCATCAACAGTGAATGTTA	180						
DB	337204	GCAAGAGTGTGGAGAAATTTTTTCAGAAATGCTGCACCATCAACAGTGAATGTTA	337145						
QY	181	GAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGC	240						
DB	337144	GAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGC	336065						
QY	241	CAAAACCAAAATATCGGGTCTGAGATCTCAAGCTGCATTTGCTGTGGGATCGGACT	300						
DB	336064	CAAAACCAAAATATCGGGTCTGAGATCTCAAGCTGCATTTGCTGTGGGATCGGACT	336305						

Db 37024 GGGGCTCCTGTTCTTTTGGGATTGTCTTAAGCCATGGGTTGGTACATCTCTGTTGC 36965  
 Qy 361 AGTTTATATGCTGAGCAACATGGAAGTTGGTAGCAGATTAAAGAGCGGCTATCA 420  
 Db 36964 AGTTTATATGCTGAGCAACATGGAAGTTGGTAGCAGATTAAAGAGCGGCTATCA 36905  
 Qy 421 TTTGCACTGCGGATTATGATCATCGCGGTCAGGCAATCATCTGGTGTCTATGCCCAT 480  
 Db 36904 TTTGCACTGCGGATTATGATCATCGCGGTCAGGCAATCATCTGGTGTCTATGCCCAT 36845  
 Qy 481 TGGCAGATGGGCAATTTGGCGCTTTTGTGGCCACTGTGTGGTGTCTATGGTATTTCCG 540  
 Db 36844 TGGCAGATGGGCAATTTGGCGCTTTTGTGGCCACTGTGTGGTGTCTATGGTATTTCCG 36785  
 Qy 541 AATTTTCTACATGCGCGGAAAGAGCCGCAACTATTATGGAGACACCTCTGTGGG 600  
 Db 36784 AATTTTCTACATGCGCGGAAAGAGCCGCAACTATTATGGAGACACCTCTGTGGG 36725  
 Qy 601 CATCTTGTGCTCACTCGATTCCATTGTCGAGCTTCGCTCGATGCTGCTGAT 660  
 Db 36724 CATCTTGTGCTCACTCGATTCCATTGTCGAGCTTCGCTCGATGCTGCTGAT 36665  
 Qy 661 GCARACAAATCCATCCCGGTACATATTTTCAATTTGAGCTTCATGCTGTGTGATCGC 720  
 Db 36664 GCARACAAATCCATCCCGGTACATATTTTCAATTTGAGCTTCATGCTGTGTGATCGC 36605  
 Qy 721 ATCGGATGCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCGCGTT 780  
 Db 36604 ATCGGATGCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCGCGTT 36545  
 Qy 781 GGTGAGTCGAAGAGCTCTTGGAGAGCTTTGGCGGCTCCATTGCTTAGGATCGGTAC 840  
 Db 36544 GGTGAGTCGAAGAGCTCTTGGAGAGCTTTGGCGGCTCCATTGCTTAGGATCGGTAC 36485  
 Qy 841 TGGTCACTCACTGTTCACTTCTCTGCTGATCACCACCTGGTGGATGGTGTGATCTTGGG 900  
 Db 36484 TGGTCACTCACTGTTCACTTCTCTGCTGATCACCACCTGGTGGATGGTGTGATCTTGGG 36425  
 Qy 901 TTGTCCTCAGTGTGTCGCGCACTGGTGGTGTGATCTGAGTTCGAGTTCAACCGCA 960  
 Db 36424 TTGTCCTCAGTGTGTCGCGCACTGGTGGTGTGATCTGAGTTCGAGTTCAACCGCA 36365  
 Qy 961 TTGGSATCAAGGATATGTCGACCTTCTTCAGGCGCACCGCGATTGATGACCGTT 1020  
 Db 36364 TTGGSATCAAGGATATGTCGACCTTCTTCAGGCGCACCGCGATTGATGACCGTT 36305  
 Qy 1021 GGATGGCATGCTCCCGCGCGATGGTGAAGTGTGATCTGAGTGTGATCAGCAGCTC 1080  
 Db 36304 GGATGGCATGCTCCCGCGCGATGGTGAAGTGTGATCTGAGTGTGATCAGCAGCTC 36245  
 Qy 1081 GTATCCGTCGTAAGCTTGGCGGAGCTTTAGTTTCAAAACTTTGAAGCGCTGAGTG 1140  
 Db 36244 GTATCCGTCGTAAGCTTGGCGGAGCTTTAGTTTCAAAACTTTGAAGCGCTGAGTG 36185  
 Qy 1141 CATAACGTGTGACCTCAGCGCTTTTGGTGTCAAAGTTTAAAGGCGCTTTACGGATT 1200  
 Db 36184 CATAACGTGTGACCTCAGCGCTTTTGGTGTCAAAGTTTAAAGGCGCTTTACGGATT 36125  
 Qy 1201 TTTTAACTGGGCGGCTGATCAAACTGCGCAGCCACCAAGCGGCAATCAATG 1260  
 Db 36124 TTTTAACTGGGCGGCTGATCAAACTGCGCAGCCACCAAGCGGCAATCAATG 36065  
 Qy 1261 CACCGTAATGGCTGCTAGTAGGAACCGATTTCGGGTGG 1300  
 Db 36064 CACCGTAATGGCTGCTAGTAGGAACCGATTTCGGGTGG 36025

RESULT 5  
 CGL312009/c 1194 bp DNA linear BCT 30-JAN-2002  
 LOCUS Corynebacterium glutamicum partial cdsa gene for CDP-diacylglycerol  
 DEFINITION synthesis.  
 ACCESSION AJ312009

VERSION AJ312009.1 GI:16505727  
 KEYWORDS CDP-diacylglycerol synthase; cdsa gene.  
 SOURCE Corynebacterium glutamicum  
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 REFERENCE 1 Nampootheri, K.M., Hoischen, C., Bathe, B., Mockel, B., Pfeifferle, W., Krumbach, K., Sahm, H. and Eggeling, L.  
 AUTHORS Expression of genes of lipid synthesis and altered lipid composition modulates L-glutamate efflux of Corynebacterium glutamicum  
 TITLE Appl. Microbiol. Biotechnol. 58 (1), 89-96 (2002)  
 JOURNAL 21820045  
 MEDLINE 11831479  
 PUBMED 2 (bases 1 to 1194)  
 REFERENCE Eggeling, L.  
 AUTHORS Direct Submission  
 TITLE Submitted (25-APR-2001) Eggeling L., Biotechnology,  
 JOURNAL Forschungszentrum, Forschungszentrum, D-52425 Juelich, GERMANY  
 FEATURES Location/Qualifiers  
 1..1194  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1718"  
 1..1194  
 /gene="cdsa"  
 <1..>1194  
 /gene="cdsa"  
 /product="CDP-diacylglycerol synthase"

ORIGIN  
 Query Match 91.8%; Score 1193; DB 1; Length 1194;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GAAGTCGTGCTCGCAAGAAAGAACTGATGGAGGCTTAGACACCTTTATCGCAATG 60  
 Db 1193 GAAGTCGTGCTCGCAAGAAAGAACTGATGGAGGCTTAGACACCTTTATCGCAATG 1134  
 Qy 61 GCTTAAGAAATACGTCGTTTCAACGTCGATTTGGCGGGGAAACGACGCTTTCTTTGCTT 120  
 Db 1133 GCTTAAGAAATACGTCGTTTCAACGTCGATTTGGCGGGGAAACGACGCTTTCTTTGCTT 1074  
 Qy 121 GCAAGAGTGTTCGAGAAATTTTTCGAAATGCTGGCACCACATCAACAGTGACATTGTTA 180  
 Db 1073 GCAAGAGTGTTCGAGAAATTTTTCGAAATGCTGGCACCACATCAACAGTGACATTGTTA 1014  
 Qy 181 GAAACTTCAAGGAGAAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 240  
 Db 1013 GAAACTTCAAGGAGAAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 954  
 Qy 241 CAACCCCAAAATAATGCGGTCGAGATCTCAAGCTGCCATTGCTGTGGGATCGGACT 300  
 Db 953 CAACCCCAAAATAATGCGGTCGAGATCTCAAGCTGCCATTGCTGTGGGATCGGACT 894  
 Qy 301 GGGGTCCTGGTCTTTTGGGATTTGCTTAAGCCCATGGGTTGTTACATCTCGTTGC 360  
 Db 893 GGGGTCCTGGTCTTTTGGGATTTGCTTAAGCCCATGGGTTGTTACATCTCGTTGC 834  
 Qy 361 AGTGTATGCTGCGCAACATGGAAGTTGGTAGCAGATTAAAGAGCGGCTATCA 420  
 Db 833 AGTGTATGCTGCGCAACATGGAAGTTGGTAGCAGATTAAAGAGCGGCTATCA 774  
 Qy 421 TTTGCCACTGCCGATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCATT 480  
 Db 773 TTTGCCACTGCCGATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCATT 714  
 Qy 481 TGGCAGATGGGCAATTTGGGCTCTTTTGTGGCCACTGTGTGGTGTGATTTTCG 540  
 Db 713 TGGCAGATGGGCAATTTGGGCTCTTTTGTGGCCACTGTGTGGTGTGATTTTCG 654  
 Qy 541 AATTTTCTACATGCGCAAGAAAGAGCCGCAACTATTGGAGGACACCTCTGTGGG 600

/translation="WPKPKNAGROLKAAIANGVGLGVLLGLVILSPGWGNYILVAGF  
MAAATWEVSRLKEGYLPLIPMTILGGQALVSPFPGTGNNAFVATVILVLMVCR  
IFYNGTEKZARNYLDTSGYIFVLMTILPGFSFAMLSLGMNNISIPGTYILVFMLCV  
IASDVGYYIAGVPFGSHPNAPLVPSPKSWEGFAGSIVLGSVTGALSVEHFLDDHHHWGIG  
LVLGLVLCVATLGLDVEISQFKRDLIGKDMENLILFGHGGLMDRLDGLMPLAANYVTLLIL  
SVTSSGVPSS"

Query Match	Score 978;	DB 6;	Length 978;
75.2%			

```
Query Match          75.2%; Score 978; DB 6; Length 978;
Best Local Similarity 100.0%; Pred. NO. 1.2e-295;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 196 ACCCATGAATGAACCGGAGCAACATACCCGGTCCATCAGGATGCCCAACCCCAAAATAA 255

Db	61	ACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCCCAACCCCAAAAAATAA	120
Qy	256	TGCGGGTCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGACATGGGGTCTCTGGTTCT	315

Db 121 TCGGGTTCGAGATCTCAAAGCTGCGCAATGCTGTGGGATCGAATCGGGGTCTCTGGTTCT 180

QY  
316  
181  
Db

Qy	376	AGCAACATGGGAAGTTCGTAGCAGACTTAAAGAGGCGGTATCATTTGCCACTGCCGAT	435
Dp	241	AGCAACATGGGAAGTTCGTAGCAGACTTAAAGAGGCGGTATCATTTGCCACTGCCGAT	300

QY 436 TATGATCATCGGGGTCAGGGAATCATCTGGCTGTCTATGGCCATTTGGCAGATGGGCAT 495

DB	301	TATGATCATCGGGCGGTGAGGCAATCATCTGGTGTCTATCGGCCATTTTGGCACGATGGGCAT	369
QY	496	TTTGGGCGTCTTTTGTGGCCACATGCTGTTGGTGTGATGATATTTCCGGAATTTTCTACAATGG	555

Db 361 TTTGGCGCTTTTGTGGCCACTGTTGGTGTGATGTAATTCGGAAATTTCTACAATGG 420

QY 55% CACGGAAAAAGAGCCCGCCACCTCTGTGGGCATCTTCGTGCTCAC 480

Qy	61.6	CTGGATTCCATTGTTTCGGAAGCTTTCGCTGCGATGCTGTCGCTGATGCAAAACAATTTCCAT	675
Dβ	481	CTCGATTCCATTGTTTCGGAAGCTTTCGCTGCGATGCTGTCGCTGATGCAAAACAATTTCCAT	540

QY 676 C C C G G G T A C A T A T T C A T T T T G A C G T T C A T G C T G T G T G A T C G C A T C G G A T G G G C G 735

DB	541	CCCCGGTACATATTTCATTTTTCACGTTTCATGCTGTGTGTGATCGCATCGGATGTGGGCGG	601
QY	736	GNATATCCGGGGTGTGTTCTTTTCGATCCGACCCCAATGGCGCCCTTGGTTCAGTCCGAAGAA	795

DB 601 GTATATCGCGGGTGTGTTCTTTGGATCGCACCCCAATGGCGCCGTTGGTAGTCCGAAGAA 660

QY 796 GTCTGGGAAGGCTTGTGCGGCTCCATTGCTTAGGATCGGTCACTGGTGACAC\*GAGTGT 720

	856	TCATTCTCGCTGGATCAACCACTGGGTGAGTGGGTGGTGCCTTAGTTGT	915
Qy			
	721	TCATTCTCGCTGGATCACACACTGGTGATGGGTGGTGCCTTAGTTGT	780
pB			

QY 9-16 GTGCGCCACGTTGGGTGACTTGGTTGAGTCGCAGTTCACACGCATTTGGGCATCAAGGA 975

Db	781	GTGCGCCACGTTGGTGACTTTGGTTGAGTCGCAGTTTCAAACGCGCATTTTGGCGCATCAAGGA	84
Qy	976	TAATGTGGAACCTTCTTCCAGGCCACGGCCGATTCATGACCCGTTTGGATGGCATGCTCCC	103

Db 841 TATCTCGAACCTTCTCCAGGCCACGGCGATTGATGGACCGTTTGGATGGCATCTCCC 900

1036 GGCCTCGATGGTGCAGCTGGTGGATCCCTGAGCTGATGATCAGCAGCCTCGTATCCCTGCTATTTAT 1037

```

RESULT 6
AX064835
LOCUS
DEFINITION
Sequence 1117 from Patent WO0100843.
ACCESSION
AX064835
VERSION
AX064835.1 GI:12542547
KEYWORDS
Corynebacterium glutamicum
SOURCE
Corynebacterium glutamicum
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1
AUTHORS
Pompejus,M., Kroeger,B., Schroeder,H., Zeider,O. and Haberhauer,G.
TITLE
corynebacterium glutamicum genes encoding metabolic pathway
JOURNAL
Patent: WO 0100843-A 1117 04-JAN-2001;
BASF AKTIENGESSELLSCHAFT (DE)
FEATURES
Location/Qualifiers
1..978
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..958
/note="unnamed protein product; RYA01894"
/codon_start=1
/transl_table=11
/protein_id="CAC25657.1"
/protein_accession="GI:12542548"
/db_xref="REFSEQ:R01894"
/db_xref="EMBL:R01894"
CDS

```

Db	901	GCCCGGATGGAGCTGGTTGATCCTGAGTGTGATCAGCAGCTCATTCCTGCTGAAG	960
Qy	1096	CTTGGGCCAGCTTTAAGT	1113
Db	961	CTTGGGCCAGCTTTAAGT	978
RESULT 7			
AX066517			
LOCUS	AX066517	978 bp	DNA linear PAT 24-JAN-2001
DEFINITION	Sequence 99 from Patent WO0100805.		
ACCESSION	AX066517		
VERSION	AX066517.1	GI:12544225	
KEYWORDS	Corynebacterium glutamicum		
SOURCE	Corynebacterium glutamicum		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
REFERENCE	Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
AUTHORS	Pompejus, M., Kroeger, B., Schroeder H., Zelder O. and Haberhauser, G.		
TITLE	corynebacterium glutamicum genes encoding proteins involved in membrane synthesis and membrane transport		
JOURNAL	Patent: WO 0100805-A 99 04-JAN-2001;		
FEATURES	BASE AKTIENGESSELLSCHAFT (DE)		
source	Location/Qualifiers		
	1..978		
	/organism="Corynebacterium glutamicum"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1718"		
CDS	101..958		
	/note="unnamed protein product; RXA01894"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAC26486.1"		
	/db_xref="GI:12544226"		
	/db_xref="REMBL:CAC26486"		
	/translations="MPKNNAGRDLKAIAVIGLGVLGLVILSPGWYILVAF MAAAETVEGRRLKEGGTHLPIMIIIGGOAILWLSWPFMTGSLASFAVTVLVLMYER IFYNGTEKEARNLRDTSVGIFVLTPILFGFPAALSLMGNNSIPGTYFLITPLMCV IASDVGYIAGVFPGSHNAPILVPSKKSWEGPAGSVILGSVTGALSIVHLIDHHWMG VILGCALVVCATLGDIVSQFKRDLDGIKDMSNLPLPHGGLMDRLDGMLPAANTVLLI SVISSVYP"		
ORIGIN			
Query Match	75.2%; Score 978; DB 6; Length 978;		
Best Local Similarity	100.0%; Pred. No. 1.2e-295;		
Matches 978; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	136	AGAATTTTTTCGAAAATGCTGGCACCATCAACAGTGACATTCTTAGAAACTTCAAGGAGA	195
Ddb	1	AGATTTTTTCGAAAATGCTGGCACCATCAACAGTGACATTCTTAGAAACTTCAAGGAGA	60
Qy	196	ACCATGAAATGAACCGGAGCAACATCACCGGTCCATAGGATGCCAACCCAAAAATAA	255
Ddb	61	ACCCATGAATGAACCGGAGCAACATCACCGGTCCATAGGATGCCAACCCAAAAATAA	120
Qy	256	TGCGGGTTCGAGATCTCAAAGCTCCCAATTCGTGTGGGGATCGAGCTGGGGGTCTCGTTCT	315
Ddb	121	TGCGGGTTCGAGATCTCAAAGCTCCCAATTCGTGTGGGGATCGAGCTGGGGGTCTCGTTCT	180
Qy	316	TTTGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTGTCAGGTTTATGGGTGC	375
Ddb	181	TTTGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTGTCAGGTTTATGGGTGC	240
Qy	376	AGCAACAATGGAAGTTGGTAGCAGACTTAAAGAAGCGCGTATCATTTGCCACTGCCGAT	435
Ddb	241	AGCAACAATGGAAGTTGGTAGCAGACTTAAAGAAGCGCGTATCATTTGCCACTGCCGAT	300
Qy	436	TATGATCATCGCGGTTCAGCAATCATCTGGCTGTGTCATGGCCATTGGCAGATGGGCAT	495
Ddb	301	TATGATCATCGCGGTTCAGCAATCATCTGGCTGTGTCATGGCCATTGGCAGATGGGCAT	360

```
QY 236 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGCTGGGGATC 295
Db 1 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGCTGGGGATC 60
QY 296 GGACTGGGGGTCCTGGTTCTTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 355
Db 61 GGACTGGGGGTCCTGGTTCTTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 120
QY 356 GTTCAGGTTTATGGCTGCGAGCAACATGGGAAGTTGGTAGCAGACTAAAGAGCGGC 415
Db 121 GTTCAGGTTTATGGCTGCGAGCAACATGGGAAGTTGGTAGCAGACTAAAGAGCGGC 180
QY 416 TATCATTTGGCACTGCCGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 475
Db 181 TATCATTTGGCACTGCCGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 240
QY 476 CCATTTGGCAGATGGGCAATTTGGCGTCTTTTGGCCACTGTGTTGGTCTGATGTAT 535
Db 241 CCATTTGGCAGATGGGCAATTTGGCGTCTTTTGGCCACTGTGTTGGTCTGATGTAT 300
QY 536 TTCCGAATTTTCTACAATGGCAGCAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 595
Db 301 TTCCGAATTTTCTACAATGGCAGCAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 360
QY 596 GTGGGCATCTTCGTGCTCACTCGGATTCATTTGGAAGCTTCGCTGGAGTCTGTGCG 655
Db 361 GTGGGCATCTTCGTGCTCACTCGGATTCATTTGGAAGCTTCGCTGGAGTCTGTGCG 420
QY 715 CTGATGCAAAACAATTCATCCCGGTACATATTTGAGTTTCACTGTTGAGTTTCACTGTTG 715
Db 421 CTGATGCAAAACAATTCATCCCGGTACATATTTGAGTTTCACTGTTGAGTTTCACTGTTG 480
QY 716 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTATATCGCGGGTGT 775
Db 481 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTATATCGCGGGTGT 540
QY 776 CCGTTGTTGAGTCCGAAGAAGCTTTGGGAAGCTTTGCGCGCTCCATGCTTAGGATCG 835
Db 541 CCGTTGTTGAGTCCGAAGAAGCTTTGGGAAGCTTTGCGCGCTCCATGCTTAGGATCG 600
QY 836 GTCACTGTTGACACTCAGTTGTTCACTTCTGCTCGATCACCACTGGTGGATGGGTTGATC 895
Db 601 GTCACTGTTGACACTCAGTTGTTCACTTCTGCTCGATCACCACTGGTGGATGGGTTGATC 660
QY 896 TTGGGTTTGCCCTAGTTGTTGCGCCACGTTGGGTGACTGTTGTTGAGTCCAGTTCAAA 955
Db 661 TTGGGTTTGCCCTAGTTGTTGCGCCACGTTGGGTGACTGTTGTTGAGTCCAGTTCAAA 720
QY 956 CGCGATTTGGGCATCAGGATATGCGAACCTTCTTCAGGCCACGCGGATTCATGGAC 1015
Db 721 CGCGATTTGGGCATCAGGATATGCGAACCTTCTTCAGGCCACGCGGATTCATGGAC 780
QY 1016 CGTTTGGATGGCATGCTCCCGCCCGCATGGTGTGACGTGGTTGATCCTGAGTTGATCAGC 1075
Db 781 CGTTTGGATGGCATGCTCCCGCCCGCATGGTGTGACGTGGTTGATCCTGAGTTGATCAGC 840
QY 1076 AGCTCGTATCCGTCG 1090
Db 841 AGCTCGTATCCGTCG 855
RESULT 9
BD16417
LOCUS BD16417 855 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD16417
VERSION BD16417.1 GI:27870229
KEYWORDS JP 2002191370-A/2216.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 855)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
```

```
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
Novel polynucleotide
Patent: JP 2002191370-A 2216 09-JUL-2002;
XJOWA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/2216
PD 09-JUL-2002
PP 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P1/40, C12P13/
PC 04, C12P13/08,
PC C12P15/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, (C12N1/21, C12R1:15),
PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
PH Key Location/Qualifiers
FT source 1. .855 /organism='Corynebacterium glutamicum'.
FEATURES
source
1. .855
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 65.8%; Score 855; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 5.8e-257;
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGCTGGGGATC 295
Db 1 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGCTGGGGATC 60
QY 296 GGACTGGGGTCTCGTCTCTTTTGGGGATTGCTTAAGCCCATGGGTTGGTATCATCTC 355
Db 61 GGACTGGGGTCTCGTCTCTTTTGGGGATTGCTTAAGCCCATGGGTTGGTATCATCTC 120
QY 356 GTTGCAGGTTTATGCTCGACCAACATCGGAAGTTGGTAGCAGACTTAAAGAGCGGC 415
Db 121 GTTGCAGGTTTATGCTCGACCAACATCGGAAGTTGGTAGCAGACTTAAAGAGCGGC 180
QY 416 TATCATTTGGCACTGCCGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 475
Db 181 TATCATTTGGCACTGCCGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 240
QY 476 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGCCACTGTGTTGGTCTGATGTAT 535
Db 241 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGCCACTGTGTTGGTCTGATGTAT 300
QY 536 TTCCGAATTTTCTACAATGGCAGCAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 595
Db 301 TTCCGAATTTTCTACAATGGCAGCAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 360
QY 596 GTGGGCATCTTCGTGCTCACTCGGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 475
Db 361 GTGGGCATCTTCGTGCTCACTCGGATTTATGATCATCGCGGGTCAGGCAATCATCTGCTGTATGG 420
QY 656 CTGATGCAAAACAATTTCCATCCCGGGTACATATTTGAGTTTCACTGTTGAGTTGTTG 715
Db 421 CTGATGCAAAACAATTTCCATCCCGGGTACATATTTGAGTTTCACTGTTGAGTTGTTG 480
QY 716 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTATATCGCGGGTGT 775
Db 481 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTATATCGCGGGTGT 540
```



```
QY 776 CCGTGTGGTGGAGTCCGAAGAGCTTTGGGAAGGCTTTGCCGGTCCATCTGTCTTAGGATCG 835
Db 541 CCGTGTGGTGGAGTCCGAAGAGCTTTGGGAAGGCTTTGCCGGTCCATCTGTCTTAGGATCG 600
QY 836 GTCACTGGTGGCACTCAGTGTCTTCACTTCTCGTTCGATCAACCACTGGTGGATGGGTGTGATC 895
Db 601 GTCACTGGTGGCACTCAGTGTCTTCACTTCTCGTTCGATCAACCACTGGTGGATGGGTGTGATC 660
QY 896 TTGGCTTGGCCCTAGTGTGTGGCCACAGTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 955
Db 661 TTGGGTGTGGCCCTAGTGTGTGGCCACAGTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
QY 956 CCGCATTTGGGCATCAAGGATATGTGCAACCTCTTCCAGGCCACGCGGATGATGATGAC 1015
Db 721 CCGCATTTGGGCATCAAGGATATGTGCAACCTCTTCCAGGCCACGCGGATGATGATGAC 780
QY 1016 CTTTGGATGGATGCTCCCGCCCGGATGTGACGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1075
Db 781 CTTTGGATGGATGCTCCCGCCCGGATGTGACGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840
QY 1076 AGCTGTATCCGTCG 1090
Db 841 AGCTGTATCCGTCG 855

RESULT 10
AP005220/c
LOCUS
DEFINITION
  Corynebacterium efficiens YS-314 DNA, complete genome, section
  7/11.
AP005220 BA000035
AP005220.1 GI:23493593
  Corynebacterium efficiens YS-314
  Corynebacterium efficiens YS-314
  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
  Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
  1 Nishio,Y., Nakamura,Y., Kawarabayasi,Y., Usuda,Y., Kimura,E.,
    Sugimoto,S., Matsui,K., Yamagishi,A., Kikuchi,H., Ikeo,K. and
    Gojobori,T.
  Comparative Complete Genome Sequence Analysis of the Amino Acid
  Replacements Responsible for the Thermostability of Corynebacterium
  efficiens
  Genome Res. 13 (7), 1572-1579 (2003)
  2 (bases 1 to 303250)
  12840036
  Kawarabayasi,Y., Yamazaki,J., Hino,Y., Kikuchi,H. and
  Director-General of Biotechnology Center.
  Direct Submission
  Submitted (17-MAY-2002) Director-General of Biotechnology Center,
  National Institute of Technology and Evaluation, Biotechnology
  Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan
  (E-mail:bicemite.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
  Kowarabayasi, Y. is officially affiliated with the National
  Institute of Advanced Industrial Science and Technology, Tsukuba,
  Ibaraki, 305-8566 Japan
  Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the
  National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
  Itoh, T. is at the Japan Biological Information Research Center,
  Koto-ku, Tokyo, 135-0064 Japan
  Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,
  Hachioji, Tokyo, 192-0392 Japan
  Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,
  Inc., Kawasaki, Kanagawa, 210-8581 Japan
  The other authors are at the National Institute of Technology and
  Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
  Location/Qualifiers
  1..303250
  /organism="Corynebacterium efficiens YS-314"
  /mol_type="genomic DNA"
  /strains="YS-314"
```

```
CDS
/db_xref="taxon:196164"
60..941
/note="CB1753, similar to AE007099-6|AAK46964.1| percent
identity: 46 in 283 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18563.1"
/db_xref="GI:23493594"
/transl_table=11
/translation="MTFRGDISRSGKARKGGGGMVGGGLGGLLVLGLVLLGGN
PAQVELLQGGDQQLFETGSCQEDANTYDDCLYTFMFSVNEVMEQVLPQ
AEQVSEVPELFTNTTSGCCASAAATPFYPADQYAYFDLTFDQMRQFGGAP
FAQMYTVAEVEHGVQNLGTLGSDYNDPGASNAVBLOADCYAGWASADKGE
DALLFVPEEELDAALTAAAVGDDNIQRRSGGEVNPESWTHSGSEQRQAPLDGYTT
GOMSSCDYLDRGVYRTT"
complement(938..3598)
/note="CB1754, similar to AB056583-2|BAB64333.1| percent
identity: 34 in 708 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18564.1"
/db_xref="GI:23493595"
/transl_table=11
/translation="NDRTHLOTQTAQOHLNRAQAOEVLRLSPARVTHRKIGF
GEPQSRVILRDADLDSANFYTRVALDGAERQLRYLTATAVLPQMETARAGT
GGFLRRFPESIRRGAEALADLTTRVADLDKQFPGLTPLDITATTATQGGVHLF
PGAHVPGQRYIDARAEIARQALGEDAPGFOWDGHHLASVHRLVHTYASDNPQVLLR
AAERALGICHTQRVEILLEQLPVDALRTATNRLRFGSLDTIGTSTVADVLRASESL
LATAQSGHOTARHKMAAQTLQOALGRNTVLIGDTPQAAVALIRILAEFEQADAL
SPAERHRRNVDTYQOLPAISATAGTVVVGAGPDGQDAWNRIITDGLAWVSANPN
LFHPQHIPTPADIDDDVLTFRAYHOGLLATLLGREIEGVDEHLDITVLQIRLELFL
TTHLDLHURGQSGFARFTVVKVLLQDDMGKTIQATAVAHRLTAEDTFTLV
VPLASVIVNWESEHRTDLPVFAHQDQKAAVTAWADTNGILCTTAGTRSLDIPA
VGLIIDEAHMKNPQSRKQCARLIDAADYALMTGTPLENKYSEFVALIRYVQPE
LITRGSAMSGDFRARIAPAYLRNQTDLDELPERDSDIMELTADPADVADQV
RAGNMWGRSAMLAKPDSATKMKRIIELLDEAEHGRKTLFTFLDLIDELHVV
LCARVTIGDGSVPARQELVDALADAPGASALISQITAGVGLNIOASQVILCEP
QVKPLIEQAVARVHRMQGTATVYHRLIGDGTADERMLEILAGTKQIFDYVARLSDI
AGTPDAVDITEAQLAASIIDAEARLGLD"
3775..4158
/note="CB1755"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC18565.1"
/db_xref="GI:23493596"
/transl_table=11
/translation="MPALDPHAPPELKHQKREATMKDTHGCHDDAPPFRPRRESYR
SRDIQVDITEDEGDSLELDGTFVEVAYVNGKYSTAHMYVEFDLTDEVVDQIRTS
DEYWLHPPEGDGPNPHRSPGTGE"
4162..5409
/note="CB1756, similar to M11302-1|AAA26834.1| percent
identity: 45 in 295 aa"
/codon_start=1
/transl_table=11
/product="putative tyrosinase"
/protein_id="BAC18566.1"
/db_xref="GI:23493597"
/transl_table=11
/translation="WHIRQNIYALTDEQLERFQDAVNGIKADGYTHFTEQHHSMEH
ATVFWESGGLHRLNSAHRGAPLFWHYCYREFELALQEKDPSVPLPYWMAWDSAA
GAAPLWNTDPDRRIYIGNGTGDSVTNTGPFAGVSVLVWGMGFRFRGGLRQL
GLMPTLPTFDEVTACITDLPVYDTPWHPGSADSFNQLGPNWGPNGHNRVHVWG
GVAGTGPNDVDFYLFHAFVLDIWARQOQTHAGGLPDGGGPPGHNLHDEMYLDST
SPTPANCIDYRHSMGYIYDPTPTQCMTSALDFTNVNAGDVERVYALFHARAQOVT
FVPLTRGLQAPYSLFTPEPTIIEHPVVDNRPYDEVRIWFATADLTGCPAPAGSAQV
RCRETDEIFEISLTGSTRELSVR"
5411..5829
/note="CB1757, similar to AE007745-6|AAK80407.1| percent
identity: 22 in 221 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18567.1"
/db_xref="GI:23493598"
```



/translation="MTPTPTPRGASCWGARGVEKGPWLQWVDMGARDTTPADPAVA  
PLASNDIEAHLRSQLNKAKSDSFRNVRVLAERKHLISAQSEALDAFELRNAISHG  
PYNLRPIADRPPTVAMI OHIRDLDDPFALEVLHSHQVRSIQDQDPPIAALKIIR  
TEISQPVYDGSYVALLTNTIARWAADSDNSILDTASTVAEVLDAETSDSAVF  
MPRTATQAEVDVNTGPELPWSVIVTEHGKPSQKPLRLISGRDMRSLIELVELR"  
6300..7421  
/note="CE1758, similar to AL049707-13|CAB41282.1| percent  
identity: 73 in 356 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC18568.1"  
/db\_xref="GI:23493599"  
/translation="MQSGITIGDVTPTDPTNGTPTTEGERIRAMTOIALKAEEVGLDV  
PATGEHNPFVSSPTTHLAYIAAOPTNLLI STATLITNDPVIKAEYALQHLIS  
GVRDLMMGRNTGPTVIFPGKDIRQGIPLAENIHLLRLWRDENVVMGKRTPLQ  
GYSTPAPLGVAFVWHGSIIRSTEIAEAFYGDGFHNNIFWNKEHQAOMNRYLQ  
RFEYIGRADQATVGGGVFGIDTESQAKDFRPYFNAPYHGPSMEDSTRTP  
LTGTAGQVTERTEFADWVDYQORQLFLIDHAGLPLEVLDQIERLGRDVPVRRR  
MEERRPDHVPSPDPTHTLLGNPEHPHFHISPGQADAE"  
7444..8121  
/note="CE1759, similar to AL049707-12|CAB41281.1| percent  
identity: 52 in 185 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative oxidoreductase"  
/protein\_id="BAC18569.1"  
/db\_xref="GI:23493600"  
/translation="MRLLTVVTAGLSNPFSTRSVADOLTNAVEAATARGESVDIEVI  
EIRLLIFDLATFTSAGMSFPALDAKRLAHSGLLIATVPVTFASYSFKFMPFPAI  
ETPLVGLPLTIIAASAGTERHSLVDHAIKPMENYLEAVVPTGVAANTADFTAGV  
EFERRVRAAGELATLMQDYSVOGLGATQADQLSYRTGTVPGEFTSFADIL  
KXHDGGO"  
complement(8359..9042)  
/note="CE1760, similar to AE000498-8|AAC77234.1| percent  
identity: 27 in 203 aa"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC18570.1"  
/db\_xref="GI:23493601"  
/translation="MRSRLCLFVKGRVYTPADLTTLTDELIDRLDPGSLVADRGF  
YSFRWATQAGDALDWRINNNVKKPHVRILEDGTWLATITPSNGFQVAPRMVR  
VYVHIDTGDDEAGIVQDHLRLTLDDPDEVSDIEDLIAAYHRTWEIENFTDEIKTHQ  
RGREVLRSKSPSEIVROEIHGLCHYAIRTLMAATAASEHPVRSFYVYKALKLTR  
SVIQGSPPP"  
complement(9074..10327)  
/note="CE1761, similar to AF024666-31|AAG03382.1| percent  
identity: 84 in 408 aa"  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="BAC18571.1"  
/db\_xref="GI:23493602"  
/translation="MAAGPVSIDETTYLDELLAASPLMREVMLOGFINOILSTQADO  
VCGDIATVSDRVNTRNGYRHDFTRVGIIDVANGKLTGSGFFPWWILERRTRABR  
ALTIVILYKGVSTRMDYVATLGINLKSQVSDMAKDDDDWEEBFRPLDTG  
PLYVSCDALTMKVRGEGVYKTSVILLATGVNAEYRLLGMVATAESVWSWTGFR  
DLKARGNEVLVTSADHLGTHQHAIGEVLPNASQRCRTHFSKNLYGMSKWTQWPLS  
AMPTFIPOQPAQYQWQARVAVAFCEQKPDVADYLEEALDELIAFTNAPKSVWTKV  
WSNNPTELRNEIRRRTDVVGIFPNRDVAVVLGAVLAEQHDDMIQQKRYMSLTSLEQ  
TRAMTANILSDSITSEIRKDYA"  
complement(10434..11051)  
/note="CE1762"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAC18572.1"  
/db\_xref="GI:23493603"  
/translation="MPRHGRTPQDTRRLSDLSISVGLLTVFPADVVDEVVQAGRTQ  
QRSRLPSRVVYVITWALNSDGSYEDIADLTDLGSLWACRWEQVAPPSAIFQ  
ARVLGHEPMAALFARVAQPARFGWGANLAGKRLIIVADGTITLDDVPTPANVEHFR  
PQTRGQAGYPOARVAIAECGSHAILDAVIGGLGTGETTIGGVH"

CDS  
complement(11090..12379)  
/note="CE1763, similar to AL357523-32|CAB93388.1| percent  
identity: 55 in 418 aa"  
/codon\_start=1  
/transl\_table=11  
/product="putative histidyl-tRNA synthetase"  
/protein\_id="BAC18573.1"  
/db\_xref="GI:23493604"  
Query Match 45.9%; Score 596.8; DB 1; Length 303250;  
Best Local Similarity 67.7%; Pred. No. 4,2e-175;  
Matches 884; Conservative 0; Mismatches 412; Indels 10; Gaps 3;  
QY 1 GAAGTCGTTGCTCGCAAGAAAAGAACTGATGGAGGCTCTAGAACACCTTTATCGCAATG 60  
DB 173786 GAGATGTTTCCCGGAGAGAACCCGAGTTGATGGAGGCTCTAG--GCCCCATCATCCAG 173729  
QY 61 GCTTAAGAAATTACGTGTTTCAACGTGATGGCGGGGAAAACGACGCTTCTTTTTCGTT 120  
DB 173728 GAGTCGTCGTTTCGTGCG--TCACAGTCAGTGGCGGGAGCGCGCTTCTTCTGTTTCC 173671  
QY 121 GCAAGAGTCGTTGGAGAGAATTTTTCGAAAATGCTGGCACCATCAACAGTGACATTGTTA 180  
DB 173670 GTGTGCCCGGTGATACCCCTGGGATGCCGATGACCGGATCTGCGTTTTCATCAGTTT 173611  
QY 181 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCAACCGTCCATGAGGATGCC 240  
DB 173610 AAACCCGAGGAGAACCCCATGATGATGAGCGGAGACCATCAACCGACCGTGGATGCC 173551  
QY 241 CAACCCAAAATATCGGGTTCGAGATCTCAAGCTGCATTTGCTGTGGGATCGGACT 300  
DB 173550 CAACCCAAAATATGTCGGGTTCGATCTCAAGCGGCGCGTGGCGTGGGTGTCGGTCT 173491  
QY 301 GGGGTCCTCGGTTCTTTTGGGGATGTCCTTAAGCCCATGCGGGTTGGTACATCCTCGTTGC 360  
DB 173490 CGGACTCTTGGTCATGTCGGCATCTCTACGCCCTCGGGCTGGTACCTCGTTGTCGC 173431  
QY 361 AGTTTATGCTGACAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGGCGGCTATCA 420  
DB 173430 CGGCGCTGTTTGGCTGGCGACCTGGGAGGTTTGTGGCGCCCTGAAGGAAGCGGTTACCA 173371  
QY 421 TTTGCCACTGCCGATATGATCATCGCGCGTTCAGGCAATCACTTGGCTGTCTATGGCAAT 480  
DB 173370 CTTGCCCATGCCGATCATGATCATCGTGGGAGGCCATGGTGTGTTCTTCTGCGCGTT 173311  
QY 481 TGGCAGATGGGCAATTTTGGCGTCTTTTGTGCCCATGTTGTGGTCTGATGATTTTCGG 540  
DB 173310 TGAGACCTCCGGGATCCTTGGCTCCTCTATGTGGCACCGCTTGTGGTGTGATGTTCTCCG 173251  
QY 541 AATTTTCTACATGGCACCGAAAAGAACCCCGCAACTATTTGAGGGACACCTCTGTGGG 600  
DB 173250 GATCTTTCTAAGCGGCCCGGAGAGGAGCGGAGAACTACCTGCGGGATACCTCAGTGGG 173191  
QY 601 CATCTTTCGTGCTCACCTGGATTCATTTTTCGGAAGCTTCGCTGCGATGCTGTGCTGAT 660  
DB 173190 TATCTTTCGTGCTCACCTGGATTCATTTTCGCGAGCTTCGCGGCGATGTTGTGCTGAT 173131  
QY 661 GCAAAACAATTCATCCCGGTATACATTTTCAATTTTGAAGCTTCATCTGTGTGTGATCGC 720  
DB 173130 GGAGAAACAACGACGTCCTCGGGGACCTTCTTCATCTCCTCAGCTTTCATGTTGGGTGGG 173071  
QY 721 ATCGGATGTGGGGGTATATCGCGGTGTGTTTCTTGTGATCCACCCCAATGGCGCGCTT 780  
DB 173070 GTCCGATGTGGGTGTTATGACAGTGGCGGTGTTCTTCGCGGCCACCCATCGCGCGGC 173011  
QY 781 GTGAGTCCGAGAAAGTCTTTGGGAAGCTTTCGCGCTCCATTTGTAGGATCGGTAC 840  
DB 173010 GGTGAGCCGAGAAATCTCGGAGGTTTTCGCGGTTCATCTCTCGGGATCGGTAC 172951  
QY 841 TGTGCACTCAGTGTTCCTCTGCTCGATCACCACTGATGGGTGATGATCTTGG 900  
DB 172950 CGGTGCCATCACCGTGCATTTCTCTGCTGCACCACTGGTGTGCTCGGTCTATCTCTGG 172891

QY 901 TTGTGCGCCAGTGTGTGGCGCACGTTGGTGACTTGGTGTGAGTCGCGAGTTCAAACGGCA 960  
 Db 172890 TATCTGTCTGGTGTCTGGCGCACCCCTCGGTGACCTGGTGGATCCCAATCCAAACGGGA 172831  
 QY 961 TTTGGGCATCAAGGATATGTCGAACCTCTTCCAGGCCACCGCGGATGATGACCGGTTT 1020  
 Db 172830 CCTCGGCATCAAGGACATGTCGAATCTGTGCTCCCGGACACGGTGGCTGATGGACCGTCT 172771  
 QY 1021 GGATGCGATGCTCCCGCGCGCATGGTGCACGCTGTGATCTGAGTGTGATCAGCAGCTC 1080  
 Db 172770 CGACGGGATGCTGCGCCTCAGCGCATGGTCACTGGTGTGATCTCAGCGTGGTGGAGCTC 172711  
 QY 1081 GTATCCGTCGTAAGCTTGGCGGACGCTTTAAGTTCAAAAACTTGAAGAGGCGCTGAGGTG 1140  
 Db 172710 CTACGCGGGTGCACAGCGAGGTGTACCATCAGGAGGGCGCGTCCCGGGGTAGTAGTT 172651  
 QY 1141 CATACGGTGTCACTCAGCGCTT-----TTTGGCTGTCAAAGTTTAAAGGGCTTTAC 1194  
 Db 172650 CAGCCCGGTGCGGGCCCTCTTTTACTCCCTGGAGGCTGATCAGCACAGGGCCCTTCT 172591  
 QY 1195 GGATTTTCTTAACCTGGCGACGCTACTCAACATCGGCACGCCACCAACAGCGCCATAA 1254  
 Db 172590 GGATCTCTTGATCTTACGACGGTACTGACATGCGGACCCACCGACCATGGCCATGA 172531  
 QY 1255 TCAATGACCGGTATGGCTGTCTAGTAGAAACCAAGTCCCGGCTGG 1300  
 Db 172530 TCAACGCAACCGGTATGCGACTGAGCAGATAGGCGATACCCCGCGG 172485

RESULT 11  
 Bx248358/c  
 LOCUS  
 DEFINITION Corynebacterium diphtheriae gravis NCTC13129, complete genome;  
 segment 5/8.  
 ACCESSION Bx248358 BX248353  
 VERSION Bx248358.1 GI:38200215  
 KEYWORDS complete genome.  
 SOURCE Corynebacterium diphtheriae  
 ORGANISM Corynebacterium diphtheriae  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 1 (bases 1 to 348408)  
 Pallen,M., Bentley,S.D., Besra,G.S., Churcher,C., James,K.D., De  
 Zoysa,A., Chillingworth,T., Cronin,A., Dowd,L., Feltwell,T.,  
 Hamlin,N., Holroyd,S., Jagsis,K., Moule,S., Quail,M.A.,  
 Rabinowitz,E., Rutherford,K., Thomson,N.R., Unwin,L.,  
 Whitehead,S. and Barrell B.G. Parkhill,J.  
 The complete genome sequence and analysis of Corynebacterium  
 diphtheriae NCTC13129  
 Nucleic Acids Res. 31 (22), 6516-6523 (2003)  
 14602910  
 2 (bases 1 to 348408)  
 Cerdeno-Tarraga,A.M.  
 Direct Submission  
 Submitted (03-OCT-2003) Cerdeno-Tarraga A.M., submitted on behalf  
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust  
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:  
 amct@sanger.ac.uk  
 Location/Qualifiers  
 1..348408  
 /organism="Corynebacterium diphtheriae"  
 /mol\_type="genomic DNA"  
 /strain="NCTC13129"  
 /db\_xref="taxon:1717"  
 /note="biotype gravis"  
 complement(58..612)  
 /gene="apt"  
 /locus\_tag="DIP1369"  
 complement(58..612)  
 /gene="apt"  
 /locus\_tag="DIP1369"  
 /EC\_number="2.4.2.7"

gene  
 CDS

/note="Similar to Corynebacterium glutamicum adenine  
 phosphoribosyltransferase Apt SWALL:APT CORGL  
 (SWALL:O87330) (185 aa) fasta scores: E(): 4.9e-42, 66.27%  
 id in 172 aa, and to Escherichia coli adenine  
 phosphoribosyltransferase Apt or B0469 SWALL:APT ECOLI  
 (SWALL:P07672) (183 aa) fasta scores: E(): 5.7e-19, 40.85%  
 id in 164 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="adenine phosphoribosyltransferase"  
 /protein\_id="CAE49900.1"  
 /db\_xref="GI:38200216"  
 /translation="MTVPYADAREAIACKTFVPDFPVGPFEDLPVLADAAAFS  
 LVEDLAKNMLGADFGCLDARGFLGSAVAYKAGTGLAIRKKGLPPPHSEBY  
 SLVEYGTAALELEPAEGLELEDKVLVLDVLTGGTGLDAAKLEACGATVSGYAVVLE  
 VDLGGRRLNDAPLVINESATA"  
 complement(61..531)  
 /gene="apt"  
 /locus\_tag="DIP1369"  
 /note="HMPfam hit to PF00156, Phosphoribosyl transferase  
 domain"  
 complement(202..240)  
 /gene="apt"  
 /locus\_tag="DIP1369"  
 /note="ScanRegexp hit to PS00103, Purine/pyrimidine  
 phosphoribosyl transferases signature."  
 complement(668..2419)  
 /locus\_tag="DIP1370"  
 complement(668..2419)  
 /locus\_tag="DIP1370"  
 /note="Similar to Corynebacterium glutamicum  
 dipeptide-binding protein DciA SWALL:O87329  
 (EMBL:AF038651) (502 aa) fasta scores: E(): 4e-83, 48.61%  
 id in 434 aa, and to Mycobacterium tuberculosis  
 hypothetical lipoprotein Rv2585C precursor or MT2662 or  
 MTCY227.16 SWALL:YP85 MYCTU (SWALL:Q50636) (557 aa) fasta  
 scores: E(): 7.1e-30, 29.09% id in 574 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Conserved hypothetical protein"  
 /protein\_id="CAE49901.1"  
 /db\_xref="GI:38200217"  
 /translation="MCGQLDRVDSILVMSYNPSPQIGNVLVGKAKOTVIRIP  
 STNLKEVVRVGLLSVMAVALSSCAHPSSLSKSNQLYAVNDLVTNAGSTE  
 GASTNAQLLAGRLYPAPVSGPKGLINLSKASQVLPGLNKIVLTISDAVSQGL  
 EATCDXLLAFTAGTYDFDSYMLMKQIDKVEKPRDEVTVTLKENFGPRWQL  
 FPAGVILPAHAIRKMGVLTDLNTVLQSRDVALQPIAEVWNRGFSDDQDPDLQVS  
 SGPKYKVRGDRGEVILTRNPHYGGQPAEIESVWVPKGNLKEIKDTGSEIAEIVTS  
 SKEIDWLDRNSTGFDVQAQSGVLTELLANAGVFATRENROAFASCIQSSIASE  
 SSELGSVAPVPTARTIRANDPMAGHVLDLTEHLHSDVEOARLQAGOTIAGVEGPD  
 ARMEALVANIRSCFAGIIVDAHQESGIRDLRSFTTHGYQFDKIGTIDVVLKA  
 LDPATEFESVATTSNNIAATQAELESTWDFVPSIFLASQPRVVDVRRVNVNLDIL  
 AGIWNMDRWFPSKSK"  
 complement(2421..2473)  
 /note="Score 66: 22/22 (100%) matches, 0 gaps"  
 complement(2489..3601)  
 /gene="secp"  
 /locus\_tag="DIP1371"  
 complement(2489..3601)  
 /gene="secp"  
 /locus\_tag="DIP1371"  
 /note="Similar to Corynebacterium glutamicum SecF  
 SWALL:Q9AE06 (EMBL:AF038651) (403 aa) fasta scores: E():  
 4.4e-84, 62.14% id in 383 aa, and to Escherichia coli  
 protein-export membrane protein SecF or B0409 or Z0508 or  
 ECS0460 SWALL:SECF\_ECOLI (SWALL:P19674) (323 aa) fasta  
 scores: E(): 1.7e-11, 25.14% id in 338 aa"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Putative protein export membrane protein"  
 /protein\_id="CAE49902.1"  
 /db\_xref="GI:38200218"  
 /translation="MSGNSVLKSLYTGDDGLDFVGRRTWRTWTLIIIVCAFCIAIVR

```
misc_feature
/BLSDIGDFVGGTMMNPAQLETSEVADTFKDATGIEPELVQIVSGDTRILEINSE
/LSDOIQRVAFKQVQSQGQPTDAIGDSTVSESGSTTQRMILSMFVFLAA
/VFYIVTRLEKEMAVAIIVALLVLIIGVGVVALGVEFSPATIIIGLTLVAFSLYDT
/VIVFDKVRNTAGYLCNRTTYAEHANLAVNOTVMSISTTIISALPIIALIVAVWL
/MVGTUKDALVOLIGVIGSTFSSVPLATPILVSLKNRSKKTAEHNRQVEMERCAGEA
/QDDSAPEQSAESTQKRTVSGPQVTTGASWRPDRG
complement (2660..3496)
/genes="secF"
/locus tag="DIP1371"
/notes="HMPfam hit to PF02355, Protein export membrane
protein"
order (complement (3443..3508), complement (3101..3166),
complement (3014..3079), complement (2936..3001),
complement (2756..2821), complement (2675..2740))
/genes="secF"
/locus tag="DIP1371"
/notes="6 probable transmembrane helices predicted for
DIP1371 by TMHMM2.0"
complement (3602..5380)
/genes="secD"
/locus tag="DIP1372"
complement (3602..5380)
/genes="secD"
/locus tag="DIP1372"
/notes="Similar to Corynebacterium glutamicum SecD
SWALL:Q9AE07 (EMBL:AF038651) (637 aa) fasta scores: E():
1.2e-139, 67.52% id in 585 aa, and to Rhodobacter
capsulatus protein-export membrane protein SecD
SWALL:SECD RHOC (SWALL:O33517) (554 aa) fasta scores:
E(): 1e-12, 27.64% id in 539 aa"
/codon_start=1
/transl_table=11
/product="putative protein export membrane protein"
/protein_id="CAE49903.1"
/db_xref="GI:38200219"
/translacion="MPSQTKKVASNSRSPFKSLFGFTVLIAIYALVATADHSARP
KIGLDGGRTVLTPQGMPTEDOLAOAKSILENRVNGVSGASVSDGNTLVITY
PGEDTQARALGQSOLFEPVAKPADPTTKLLDAMKEMANRWVGVLTPOEANDA
LSNTAKALASGKKPDAKAPTVAKEPEKANSIESEOLRNKVLVWKDKQSDTA
TQLSASLLKDDTPDLQDTPKPLVACDPSSGQVYVLSFAPLQGETQANGK
RLTNEIDTRNPTGNGPTQMEIGFSPKSDNGDSATWAKLTQYELQKQVAILT
DSKIIISAPVQTSAPVGSATITGFSQKQAQLANNKYGALPLSPAGENGSTGGT
TVPASIGVASLAKGLVGLIIVLYALINYPFGLISFSLVASFALVFGLVL
LGRVGYSLDLAGTAGLIGIGTTADSFVVLVYERIKDEVDRGSRSSVPRGWERARQ
TIYNGAVTLGAVIIVILAVGEVKGAFILGLTIFDIIVTVFLVATPLVLASRKPW
TSSGAANGMKIISNHHGANAKGEE"
complement (5267..5380)
/genes="secD"
/locus tag="DIP1372"
/notes="Signal peptide predicted for DIP1372 by SignalP 2.0
HMM (Signal peptide probability 0.993) with cleavage site
probability 0.652 between residues 38 and 39;
signal-peptide site"
complement (3689..4513)
/genes="secD"
/locus tag="DIP1372"
/notes="HMPfam hit to PF02355, Protein export membrane
protein"
order (complement (5255..5320), complement (4109..4162),
complement (4028..4093), complement (3932..3997),
complement (3794..3850), complement (3698..3763))
/genes="secD"
/locus tag="DIP1372"
/notes="6 probable transmembrane helices predicted for
DIP1372 by TMHMM2.0"
complement (5385..5570)
/locus tag="DIP1373"
complement (5385..5570)
/locus tag="DIP1373"
/notes="No significant database matches"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
```

```
/protein_id="CAE49904.1"
/db_xref="GI:38200220"
/translacion="MAKRIPLNLDNFVKCSVSDSASVAYHSQEPKLLFDRVHFRIDGK
QREADGGVLLDHLHRS"
complement (5611..5904)
/locus tag="DIP1374"
complement (5611..5904)
/locus tag="DIP1374"
/notes="Similar to Corynebacterium glutamicum hypothetical
13.5 kDa protein SWALL:Q9AE08 (EMBL:AF038651) (121 aa)
fasta scores: E(): 7.2e-10, 34.83% id in 89 aa, and to
Mycobacterium tuberculosis hypothetical 13.0 kDa protein
Query Match 26.3%; Score 342.2; DB 1; Length 348408;
Best Local Similarity 63.1%; Pred. No. 5.5e-95;
Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
QY 237 TGCCCAAAACCCAAAATAATGCGGTCGAGATCTCAAAAGCTGCCATCTCTGFGGGAATCG 296
DB 142124 TGCCCTAAACCGAAGAAATTCGGCTGGTCGAAACCTGAAAGCGCGATCAGCGTTGGAATTG 142065
QY 297 GACTGGGGTCTCTGTTCTTTTGGGGATTGCTCTTAAGCCCATGGGGTTGGTACATCTCG 356
DB 142064 GCTTGGCGGCCCTCGTATTGCTCGCTATCTTTGTTTATTCCTTTTGGTGGTATCCATTAG 142005
QY 357 TTGCAAGTTTATGCTGCAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCT 416
DB 142004 TGGCTATCCATCTCAGTGGCACTTTGGAGGTAGAGCGCGTTTAAAGAGCGCGTT 141945
QY 417 ATCAATTGCCACTCGCGATATGATCATCGGCGGTGAGCAATCATCTGGCTGTCTATGCG 476
DB 141944 ATCTTTGCAAGTTGGGTCATGCTTATTTGGCGGCAAACTCATGTTGGGCTTAGTTGGC 141885
QY 477 CATTTGGCAGATGGGCATTTTGGCGTCTTTTGGCCACTGCTGTTGGTGTGCTGATTA 536
DB 141884 CATTTGGTCCGAAGGATAGTGGCTGTTTCTGTCGGCTGTACTCGTACGATGTTG 141825
QY 537 TCGCAATTTTCAATATGACACGGAAGAAAGAGCCGCACTATTTAGAGGACACCTCTG 596
DB 141824 GTGCGCTTTTCCATCATGCTGCTTCTATGCCACCGAAAACTATCTCGAGATACTGCG 141765
QY 597 TGGGCATCTTGCTGCTACCTGATCCATCTTCGGAAGCTTCGCTCGGATGCTGTGCG 656
DB 141764 TCGCAATTTTGTTTTAACTGATGCCCTTTTGGTAGTTTGGGCAATGCTTTGCG 141705
QY 657 TGATGCAAAACAATTCATCCCGGGTACATATTTTCAATTTTGAAGTTTCATGCTGTGTGCA 716
DB 141704 TGTTGAAACAGAAACAGCACCCGGAAGTACTTCATGCTGCACTTCAATGTTGCGTTA 141645
QY 717 TCGCATCGGATGFGGCGGGTATATCGCGGTGTGTTCTTTGATCGCACCCCAATGCGCG 776
DB 141644 TCGCCTCTGACGTAGGCGGTATATTTGCGAGGCGTATGTTGCGATCGCACCCCATGGCAC 141585
QY 777 CGTTGGTGTCTCGAAGAAAGTCTTTGGAAGGCTTTGCGCGCTCCATGTTCTTAGGATCGG 836
DB 141584 CTGCGGTAGCCCCAAAGAAATCGTGGAGGGCTTTATCGGTTCTGTAGTCTTGGCATGA 141525
QY 837 TCACTGGTGCACCTCAGTGTTCATCTCCTGCTCGATCACCATCTGGTGGATGGGTGTGATCT 896
DB 141524 TTGTTGGAGCATTCACCTGTGCTCTATTTATTAGGGCATCAATGCTGTGGGATTAAGTTTC 141465
QY 897 TGGGTTGTGCCCTAGTGTGTGCGCCACGTTGGGTGACTTGGTTGAGTGCACGTTCAAAAC 956
DB 141464 TCGGATTTGGACTGTGTGATCTGCGGACACTGGGCGCATCTTGTGCAATCCAGTTTAA 141405
QY 957 GCGATTTGGGCATCAAGGATATCTCGAAACCTTCTCCAGGCCACGGCGGATTCGATGACC 1016
DB 141404 GGGAACTCGGRTCAAGACATCTCTCGGATTTTGGCGGCCATGTTGGTCTGATGATC 141345
QY 1017 GTTTGGATGCGATGCTCCCGGCGCGATGTTGAGTGGTTGATCCTCTGAGTGTGAT 1071
DB 141344 GGCTTGACGGTATGTTGCCATCGGCTATGTTGATGATGGTGTGATCTTGGTGTGAT 141290
```



YVGYGLYELRLIAGVANGPQDAVIAAAGRLGALAGWYNQHGAWFVALLVVLVVG	
FASTWFRVR	
complement (5731. .6264)	
/gene="WT2946"	
CDS	
complement (5731. .6264)	
/gene="WT2946"	
/notes="similar to SP:Q10804 PID:1403398 PID:1403398	
PID:1707631 GB:AL123456; identified by sequence	
similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="soluble secreted antigen MPT53 precursor"	
/protein_id="AAK47271.1"	
/db_xref="GI:13882722"	
/translation="MRSVMSLRVSPKAFADGIVAJAVLVMFLANTPRAVADE	
RLGFTATLSGAFDGSLOGKPAVLNFWPFCNAPSLSOVAANPAVTFVGI	
ATRADVGMSQFVSKYNLFTNLNDADGIWIRINFWQPAFVYFADGTSFVNPT	
AAMSQELSGRVAALTS"	
complement (6438. .7532)	
/gene="WT2947"	
CDS	
complement (6438. .7532)	
/gene="WT2947"	
/notes="identified by Glimmer2; putative"	
/codon_start=1	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="AAK47272.1"	
/db_xref="GI:13882723"	
/translation="MVPELWDEPRGPRPHLADLAAGASAVAEGLPAPRAKOL	
AHQYGLRLDPRMDLPAARDRIAGAFNLLTASADITCDAGOTKTLRAVNG	
TMFESVLMRYPRNTVCISQAGCGACPFCACTGGGLTRNLSTAELEOVRAAAL	
RDDFGRLSNVFMGEPFLANVARVLAQVRIETAREPFGSISARAATVSTGLAPA	
INLADARLGVTLALSADGLDRLDLPVNNRIRISEALDAARYANVTGRRVIE	
VALRDYNDOPRALDLKRLHRLVGLPLAHVLIPLNPTPGSDWDASPKVEFEVFR	
VRAKGVSVTRDTEGREISAAQGLAAVG"	
complement (7555. .8604)	
/gene="WT2948"	
CDS	
complement (7555. .8604)	
/gene="WT2948"	
/notes="similar to GB:M11330 SP:P06466 PID:1208947	
PID:145476 GB:U00096; identified by sequence similarity;	
putative"	
/codon_start=1	
/transl_table=11	
/product="phosphatidate cytidyltransferase, putative"	
/protein_id="AAK47273.1"	
/db_xref="GI:13882724"	
/translation="WRSVAOKRISTRPTNTSPKLSWLNTPKASWRSGRSATKSV	
TTNDAAGNPAEQARGAKQAPATETSPAGSDLRAAIVGLSIGLVLIAVLVFPVPM	
VAVAVATLVATHEVRRRLRAGHLIPVILLIGQAVALTWEPFGVAGLAGEGVN	
VVCMIRLFMDQSVTRPTTGGAPSPGNLYSDVSATVLAHVPLFCFSGAMLVYPNG	
SGWVFCMMIAIVSDGVYAVGVLFGRHPMPTTSPKKSWEFGAGSLVCGITATITTA	
TFLVKTPIWIGALLGVLTALGDLVESQVRDLGIKDMGRLLPQHGLMDRLDGI	
LPSAVALVILTLIP"	
complement (8498. .9055)	
/gene="WT2949"	
CDS	
complement (8498. .9055)	
Query Match	
Best Local Similarity 14.8%; Score 191.8; DB 1; Length 16952;	
Matches 456; Conservative 54.4%; Pred. No. 5.6e-48;	
257 GCGGTCGAGATCTCAAGCTCCATTCTGTGGGATCGGACTGGGGTCTCTGTTCTT 316	
8388 GCGGCGCTGATTTCGCGCAGCGATCGTGGTGGGCTTCATAGGCTAGTCTCATC 8329	
317 TTGGGATGTCTTAAGCCATCGGGTGGTACATCTCTGTTCAGGTTTATGCTGCA 376	
8328 GCGGTGCTGTTCTGTTCCGCGCTTGGGTGCACTCGTGGCGCTCGCACCTTGTC 8269	
377 GCAACATGGGAATGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436	
8268 GCTACCATGAGTGTGCGGAGGTTGCGGAAGCGGCTATCTCATCCCGTTATCCG 8209	

437	ATGATCATCGCGGTCAGGCAATCATCTGGCTGTCTATGGCCATTGGCACGATGGGCATT	496
8208	TTGCTGATTGGGGGACGCGCGGTGGCTGACCTGGCGGTCGGCGCGTTCGGCGCA	8149
497	TTGGCGCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	556
8148	TTGGCGGCTTTGTGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8089
557	ACGGAAGAAAGAGCC-----GCAACTATTGTAGGACACCC	592
8088	ACGCTGACGCGCCACACACCGCGTGCACCGTCCGCGGAAACTACTTGTCCGACGTC	8029
593	TCGTGTGGGCATCTTCTGTCTCACCTGGAATTCATTGTTCGGAAGCTTCGTCGCGTGTG	652
8028	TCGGCCACGCTTCTCTGCGCGTGTGGGTCCCATTTGTCTGTCTTTCGGGCAATGTGT	7969
653	TCGCTGTATGCAAAACAATCCATCCCGGATATATTTTTCATTTGACGTTTCAATGTGT	712
7968	GTCTACCCCGAAATGGCTCGGATGGG-----TGTTCTGATGATGATCGCG	7921
713	GTGATCGCATCGATGTGGCGGATATATCGCGGATGTTCCTTTTGGATTCGCCCAATG	772
7920	GTATCGCTTCGATGTGGCGGCTACGCGGTGGGTGTCTGTTTGGCAAGCATCCGATG	7861
773	GGCGCTTGTGTAGTCCGGAAGTCTTGGGAAGCTTTGCCGCTCCATGCTCTTAGGA	832
7860	GTTCGACGATCAGCCCGAAGAGTCTGTGGGAGGCTTTGCCGCTTGTGTGTGTGTGT	7801
833	TCGGTCACTGGTGTCTCACTAGTGTTCATCTCTGTCTGATCACCCTGTGTGTGTGTGT	892
7800	ATCACCGCAACCATCATCACCGGACTTTCCTGTGGCAAAACGCGGTGATGTGTGCA	7741
893	ATCTTGGGTGTGCCCTAGT	952
7740	CTGCTCGCGTGTCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7681
953	AAACGCGATTTGGGATCAAGGATATGTCAACCTTCTTCCAGGCCACGCGGATGTGTG	1012
7680	AAACGCGATTTGGGATCAAGGATATGTCAACCTTCTTCCAGGCCACGCGGATGTGTG	7621
1013	GACCGTTTGGATGGCATGCTCCCGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1071
7620	GACCGCGTGGACGATGCTCCCGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7562
RESULT 13		
BX248344/c	318050 bp	DNA
LOCUS		
DEFINITION		linear BCT 11-JUN-2003
		Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
		11/14.
ACCESSION	BX248344	
VERSION	BX248344.1	GI:31619628
KEYWORDS		complete genome.
SOURCE		Mycobacterium bovis subsp. bovis AF2122/97
ORGANISM		Mycobacterium bovis subsp. bovis AF2122/97
		Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
		Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
		tuberculosis complex.
REFERENCE		
AUTHORS		
		Garnier, T., Eiglmeyer, K., Camus, J.-C., Medina, N., Mansoor, H.,
		Pryor, M., Duchoy, S., Grondin, S., Lacroix, C., Monsemp, C., Simon, S.,
		Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L.,
		Wheeler, P.R., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V. and
		Hawinson, G.
TITLE		The complete genome sequence of Mycobacterium bovis
JOURNAL		Online Publication
REMARK		FNAS 10.1073/pnas.1130426100 ( Microbiology )
REFERENCE		2 (bases 1 to 318050)
AUTHORS		Garnier, T.
TITLE		Direct Submission
JOURNAL		Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
		Bacterienne Institut Pasteur 28, rue du Dr Roux 75724 PARIS cedex

DEHYDROGENASE (EC 1.2.1.39) from *Escherichia coli* strain K12 (499 aa), FASTA scores: opt: 1074, E(): 3e-59, (42.2% identity in 462 aa overlap); etc. Also similar to many *M. tuberculosis* dehydrogenases e.g. P71823 [Rv0768][MTC9369.13 (489 aa), FASTA score: (38.1% identity in 467 aa overlap)]. Contains P800687 Aldehyde dehydrogenases glutamic acid active site and P800070 Aldehyde dehydrogenases cysteine active site. BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY."

ALT

```

/codon_start=1
/transl_table=11
/product="PROBABLE ALDEHYDE DEHYDROGENASE ALDC"
/protein_id="CAD96570.1"
/db_xref="GI:31619630"
/translation="MST*QLINPATEVLSVDHTDANAVDDQVRAARAAQRWRARLA
PQRAQRGAFATAAAQVADLDEALEYVANGSHITVSAEMRAGHVRODLAFAYASPERL
SGRQIPVAGGVVDTEPMGVGVITPWNPFWIASWAIAPALAAAGNAVLPKPAELTP
LITMRGLLEAVEAGLDELLOVPGKTVTGEPVTHPDRKIVTSTGVCKRVWAG
AAAQKRVTLLELGKGSANIYTHDCDLRAATAPAGVFINAGQCCARGLVQSRVY
DFMFLSPFVAHSIVGDPDGERATETMGLVSRHARDKVAGVPPDPAPVFRGTAPAGR
FGFPFTVLTKRGDRTVITFQGVTVVLTDFEADATSLANDTAYGLSGITITDGL
SRALLARAVARVSGNLSVNHSVSFAFNPFCGFQSGVGRGLGDPALPQFTEKAWFLA
VFVFVF"

```

ALSO similar to other amidotransferases and hypothetical proteins, but shorter in N-terminus e.g. O88072|SC135.37 and O96919|YK30 MYCTUR|VC2090|MT2151|MTCV49.30 STREPTOCOCCUS PNEUMONIAE PHOTOPHETICAL 25.3 KDA PROTEIN from Streptococcus coelicolor strain ATCC 25611. FASTA scores: OPT: 683, E: 1.2e-32, (47.65% identity in 235 aa overlap); AAK79730|O97188|CACI1764 IDENTIFIED GLUTAMINE AMIDOTRANSFERASE from Clostridium acetobutylicum (241 aa), FASTA scores: opt: 458, E: 1.6e-19, (32.95% identity in 246 aa overlap); Q97QV9|SP1089 GLUTAMINE AMIDOTRANSFERASE CLASS I from Streptococcus pneumoniae (229 aa), FASTA scores: opt: 5.6e-18, (34.75% identity in 236 aa overlap); Q13431, E: 5.6e-18, (34.75% identity in 236 aa overlap); Contains three 17 aa repeats at the N-terminus very similar to those in other Mycobacterium tuberculosis proteins e.g. Q106919|YK30 MYCTUR|VC2090|MT2151|MTCV49.30 STREPTOCOCCUS PNEUMONIAE PHOTOPHETICAL 25.3 KDA PROTEIN from Streptococcus coelicolor strain ATCC 25611.

```

similar to those in other Mycobacterium tuberculosis
proteins e.g. Q10699|YJ30 MYCTU|RV2090|MT2151|MTCY49.30
PREDICTED 5'-3' EXONUCLEASE RV2090 (BC 3.1.11.-).
<codon_start=1
/transl_table=1
product="POSSIBLE AMIDOTRANSFERASE"
protein_id="CAD96571.1"
db_xref="GI:31519631"
translation="MDLSASDGGDLRPSAPRSPVSDGGDLRPSAPRSPVS
GGVGLPAPRSPRGLSGRSPVGLTALVLEQVRTGVMIDPAGYLPADYFEGITMAG
VALLPQPVSDSGVGLSDLHALVITGCDLDPAAQVSEPHATDPRFGDAWE
ALLALLRGLALGRCVGLGICRGVTHQNLVAGLTGHLQPLDLHGSHGPAAGNVTRLPVHT
ASGATGSLRFLAELIGESADPVQHQAIDQVGLGVSAVDVGDVLEALELPGDITFLVAVQ
complement(3143..4516)
gene="glnA4"
locus_tag="Mb2885c"
complement(3143..4516)
gene="glnA4"
locus_tag="Mb2885c"
EC_number="6.3.1.2"
note="Mb2885c, glnA4, len: 457 aa. Equivalent to RV2860C,
note="Mb2885c, glnA4, len: 457 aa. From Mycobacterium tuberculosis strain H37Rv,
100% identity in 457 aa overlap). Probable glnA4,
glutamine synthetase class II (BC 6.3.1.2), similar to

```

OC,  
RV,

many glutamine synthases e.g. O88070|SCI35.35c from Streptomyces coelicolor (462 aa), FASTA scores: opt: 1947, E(): 8.2e-120, (64.15% identity in 452 aa overlap); Q98H15|MLJ3074 from Rhizobium loti (Mesorhizobium loti) (465 aa), FASTA scores: opt: 1321, E(): 7.8e-79, (46.7% identity in 452 aa overlap); Q98EM0|MLJ4187 from Rhizobium loti (Mesorhizobium loti) (456 aa), FASTA scores: opt: 698, E(): 4.6e-38, (33.5% identity in 454 aa overlap); Q9CDL9|GLNA from Lactococcus lactis (subsp. lactis) (Streptococcus lactis) (446 aa), FASTA scores: opt: 633, E(): 8.2e-34, (32.45% identity in 456 aa overlap); etc. Also similar to three other potential glutamine synthases in Mycobacterium tuberculosis: Q10378|GLN2|MYCTU|GLNA2|RV2222c|MT2280|MTCY190.33c|MTCY427 (0.3c PROBABLE GLUTAMINE SYNTHETASE (446 aa), FASTA score: (31.1% identity in 453 aa overlap); RV1878|glna3 and RV2220|glna1. BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY." /codon\_start=1 /transl\_table=1 /product="PROBABLE GLUTAMINE SYNTHETASE GLNA4 (GLUTAMINE SYNTHASE) (GS-1)" /protein\_id="CAD96572.1" /db\_xref="GI:131619632" /translation="MTGPGSPPLANTLEPLVAAGVDVTVIVAFDMQGLAGKRISG RHFVDIATRGVECCSILLVDVNLNVPGYMASWDGTGDMVMPFDLSTLRILPWL PGTALVADLVWADSEVAPSPILRRQLDRKARGLVADVATELEFIVFDQFYQA WASYRGLTTPASDYNIDYALASRMEPLRLDLGNAGAGLFEAVFGECNQGQBEI GFYDEALVTCDNAHYNGKATADQKSLTFMAKYDERGNSCHIVSLRGTDGS AVFADNSGPHGMSMPFSFVAGQLATREFTLCVAPTINSYKRFADSSFAPTALAWL

Query Match 14.8%; Score 191.8; DB 1; Length 318050;  
Best Local Similarity 54.4%; Pred. No. 1.1e-47;  
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;  
QY 257 GCGGTCGAGATCCAAAGTCGTCATGCGGATCGGACTCGGGGGCTCGTCTT 316  
DB 23049 GCGGCGGTGATTCGCGGACGATCGTGTGGGCTTCCATAGGCGTATCATC 22990  
QY 317 TTGGGATGTCCTAAGCCATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 376  
DB 22989 GCGGTGCTGTGTTGCTTCCGCGGTTGGTGGTGGTGGTGGTGGTGGTGGT 22930  
QY 377 GCAACATGGGAGTTGGTGGGAGTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436  
DB 22929 GCTACCCATGAGTGGTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 22870  
QY 437 ATGATCATCGCGGTGAGGCAATCATCTGCTGTGTCATGCGCATTTGGCAGTGG 496  
DB 22869 TTGCTGATTCGCGGCGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 22810  
QY 497 TTGGGCTTTTGGGCGGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 556  
DB 22809 TTGGCGGCTTTGGTGGGATGCTGTTGTCGATGATTTGGCGGACTGTTGAT 22750  
QY 557 ACGGAAAAAGAGCC-----GCAACTATTTGAGGAGACCC 592  
DB 22749 AGCGTGACGCCCGGACGACCGCGGCTGACCGTGGCGGAACTACTTTCGCGAG 22690  
QY 593 TGTGTGGGCTATTCGTGCTCACTGGATTCATTTGTTGGAAAGCTTCGCTGGCAT 652  
DB 22689 TCGGCGGCTTCTCTGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 22630  
QY 653 TCGCTGATGCAAAACAAATTCATCCCGGTGATATTTTCATTTGACGTTTCATG 712  
DB 22629 GTCTACCGGAAATGGCTCGGATGGG-----TGTTCTGATGATGATCGG 22582  
QY 713 GTGATGCGATCGGATGCGGGGTATATCGCGGTGTTGTTCTTTGGATGCCAATG 772  
DB 22581 GTGATGCTTCCGATGTCGCGGCTACCGCGTGGGGTGTGTTTGGCAATCCGATG 22522  
QY 773 GCGCGCTTGGTGGTCCGAGAGCTTTGGGAGGCTTTCGCGGCTCATTTGCTTAGGA 832  
DB 22521 GTTCCGAGATCAGCCCGAAGATCGTGGGAGGCTTTGCGGCTTCTGCTGTGCGGG 22462

QY 833 TCGGTCACTGGTGCACCTCAGTGTTCACCTTCTGCTCGATCACCACCTGGTGGATGGTGG 892  
DB 22461 ATCACGGCAACGATCATCACCGCAGCTTTCCTGTGTCGCAAAACCGCGTGGATTGGTGA 22402  
QY 893 ATCTTGGTGGTGGCCCTAGTGTGTGGCCAGTGTGGTGGTGGTGGTGGTGGTGGTGGT 952  
DB 22401 CTGCTCGGCGTCTTTTCTGTCTCACCCCGCGCTGGGCGACCTGGTGGAGTCGCGAGTTC 22342  
QY 953 AAACGCGATTTGGGATCAAGGATATGTCGAACCTTCTTCCAGGCGCAGCGCGATGATG 1012  
DB 22341 AAACGTGACTCGGATCAAGACATGGCGCGCTGTACCCGCGCACGCGCGTCTGATG 22282  
QY 1013 GACCGTTTGGATGGATGCTCCCGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071  
DB 22281 GACCGCTCGACGGCATATGCTCTTCCGCGTGGCGGCTGGATAGTCTCCTCACACTGCT 22223  
RESULT 14  
BX842581/C standard; circular genomic DNA; PRO; 348676 BP.  
ID BX842581  
XX  
AC BX842581; AL008883; AL008967; AL021070; AL021287; AL021309; AL123456;  
AC 274024; Z74697; Z81331; Z83018; Z83857; Z83858; Z83866; Z95150; Z95207;  
XX  
SV BX842581.1  
XX  
DT 21-NOV-2003 (Rel. 77, Created)  
DT 21-NOV-2003 (Rel. 77, Last updated, Version 1)  
XX  
DE Mycobacterium tuberculosis H37Rv complete genome; segment 10/13  
XX complete genome.  
XX  
OS Mycobacterium tuberculosis H37Rv  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
OC Mycobacterium tuberculosis complex; Mycobacterium tuberculosis.  
XX  
[1]  
RN MEDLINE; 98295987.  
RX PUBMED; 9634230.  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry III C.E., Tekai F., Badcock K.,  
RA Bisham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K.,  
RA Peitwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K.,  
RA Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J.,  
RA Quail M.A., Rastam M.A., Rogers J., Rutter S., Seeger K., Skelton S.,  
RA Squares S., Squires R., Sulston J.E., Taylor K., Whitehead S., Barrall B.G.,  
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete  
RL genome sequence";  
RL Nature 393:537-544(1998).  
XX  
[2]  
RN PUBMED; 12368430.  
RX Camus J.C., Pryor M.J., Medigue C., Cole S.T.;  
RA "Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv";  
RL Microbiology 148:2967-2973(2002).  
XX  
[3]  
RN 1-348676  
RP Parkhill J.;  
RA  
RT Submitted (11-JUN-1998) to the EMBL/GenBank/DBJ databases.  
RL Submitted on behalf of the Mycobacterium tuberculosis sequencing and  
RL mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton,  
RL Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut  
RL Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail:  
RL parkhill@sanger.ac.uk  
XX  
CC Notes:  
CC Details of M. tuberculosis sequencing at the Sanger Centre  
CC are available on the World Wide Web.







```

FT repeat_unit complement (3774..3809)
  Query Match 14.8%; Score 191.8; DB 15; Length 348676;
  Best Local Similarity 54.4%; Pred. No. 1.2e-47;
  Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;

Qy 257 GCGGTCGAGATCCAAAGTCGCTGTGGGATCGGACTGGGGGTCCTGGTTCIT 316
Db 72497 GCGGCGGTGATTTGCGCGAGCGATCGTGTGGGCTTCCCATAGGCTAGTCTCATC 72438

Qy 317 TTGGGGATTGCTTAAGCCCATGGGTTGGTATCATCTCTGTCAGGTTTATGGCTGCA 376
Db 72377 GCTACCATGAGTGTGGGAGGTTCGGGAACGGGCTATCTCATCCGGTATCCCG 72318

Qy 437 ATGATCATCGCGGTTCAGGCATCATCTGGCTGTCTATGGCCATTTGGCAGATGGGCATT 496
Db 72317 TTGCTGATTGGCGGCGACGGCGGTGTGGCTGACCTGGCGGTTTCGGCGGGTGGCGCA 72258

Qy 497 TTGGCGCTTTTGTGGCCACGTGTGGTGTGATGATTTCCGAATTTCTCAATGGC 556
Db 72257 TTGGCGGCTTTGTGGCATGCTGTGTCTGTCGATGATTTGGCGACTGTTCATCGAGAC 72198

Qy 557 ACGGAAAAGAGGCC-----GCAACTATTTAGGGACACC 592
Db 72197 AGCGTGACGCCCGACGACGCCGGGTGCACCGTCGCGGGAACTACTTGTCCGAGTC 72138

Qy 593 TCTGTGGCATCTTGTGCTCACTGATTCATGTTCCATGTTCCGAAGCTTCGTCGATGCTG 652
Db 72137 TCGGCCACGCTCTTCTGTCGCGGTGTGGTCCCATTTGTTCTCTTTTCGGCGCAATGCTG 72078

Qy 653 TCGCTGATGCAAAACAATTCCATCCGGGTACATATTTGACGTTTCATCTCTGT 712
Db 72077 GTCTACCGGAAATGGTCGGATGG-----TGTTCTCATGATGATCGCG 72030

Qy 713 GTGATCGCATCGGATGTTGGGGGTATATCCGGGTGTGTTCTTTGGATCGACCCAAATG 772
Db 72029 GTCATCGCTTCGGATGTCCGGGTACCCGCTGGGGGTGTGTTTGGCAAGCATCCGATG 71970

Qy 773 GCGCGTTGTTGAGTCGAGCAAGTCTTGGAGGCTTTGCGGCTCCATTGCTTAGGA 832
Db 71969 GTTCGAGATCAGCCGCAAGATCGTGGAGGCTTTGCGGTTTCTGGTGTGCGG 71910

Qy 833 TCGETCACTGTGCACTCAGTGTTCCTCTCTGATCACCACTGCTGGATGGGTGTG 892
Db 71909 ATCACCGCAACGATCATCACCGCACTTCTCTGTCGCAAAACGCGGTGATTTGTCGA 71850

Qy 893 ATCTTGGTTGTGCCCTAGTTGTGTGCCCACTGGTGTGATCTGTTGATTCGCACTTC 952
Db 71849 CTGCTCGCGCTGCTTTTCTGTGCTACCACTCGGCTGGCGACCTGGTGGATCGCAGTC 71790

Qy 953 AAACCGGATTTGGGCATCAAGGATATGTCGAACCTTCTTCCAGGCCACGGCGGATTTGATG 1012
Db 71789 AAACGTGACCTCGGATCAAGACATGGCGGCTGCTATCCGCGCACGGGCTCATG 71730

Qy 1013 GACCTTTTGGATGGATGCTCCCGCGCGGATGATGATGATGATCTGATCTGATGAT 1071
Db 71729 GACCGGCTCGAGGCATATGCTTCTCCCGGTGGCGGCTGGATAGTCTCACACTGCT 71671

RESULT 15
MSGB1529CS
LOCUS
DEFINITION
Mycobacterium leprae cosmid B1529 DNA sequence.
ACCESSION
L78824
VERSION
L78824.1 GI:1377769
KEYWORDS
Mycobacterium leprae
Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Original source text: Mycobacterium leprae (clone: cosmid B1529)
(tissue library: Lcrist 6) DNA.
This sequence was generated by the Genome Sequencing Center at
Genome Therapeutics Corporation (Collaborative Research Division),
100
Beaver St., Waltham, MA, 02154. Please contact Doug Smith
(smith@eric.com) for further information. The sequence represents
the
insert of a Lorient 6 cosmid clone from a mapped set of clones
constructed from M. leprae genomic DNA isolated from armadillo
liver
[3]. The sequence may not represent the entire cloned insert of
the
cosmid if an overlapping region was previously sequenced from
another
clone. Coding sequences larger than 60 amino acids were predicted
on
the basis of codon usage and homology information. An attempt was
made
to locate the most probable start site based on codon usage,
homology,
the presence of a Shine-Dalgarno sequence, or overlapping orf that
suggested translational coupling. It is possible that the actual
start
site differs from the one selected.
FEATURES
source
1..36985
Location/Qualifiers
/organism="Mycobacterium leprae"
/mol_type="genomic DNA"
/specific_host="Dasypus novemcinctus"
/db_xref="taxon:1769"
/clone="cosmid B1529"
/tissue_lib="Lorient 6"
/notes="The liver of the armadillo was used to isolate the
Mycobacterium leprae."
ORIGIN
Query Match 12.6%; Score 164; DB 1; Length 36985;
Best Local Similarity 52.6%; Pred. No. 3.8e-39;
Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;
Qy 257 GCGGTCGAGATCTCAAAGTCGCTGTGGGATCGGACTGGGGTCCCTGGTTCIT 316
Db 8360 GCGGGACGTAATTCGCCCGCGGATCGCGGTGGGCTTAAGTATGTTGTTCTCGTC 8419
Qy 317 TTGGGATTGCTCTAAGCCCATGGGTTGGTATCATCTCTGTCAGGTTTATGGCTGCA 376
Db 8420 GCACGCTGTGTTTGTCTCGGAACTCTGGTGTCTTGTGCGCTGGCCATTTTCGTT 8479
Qy 377 GCAACATGGAAGTTGGTAGCAGACTTAAGAGCGGCTATCATTTGCCACTGCCGATT 436
Db 8480 GCTAGCCATGAGTGTGCGCGGCTACGGAAGCTGGATATGTGATTCGGGTATCCCG 8539
Qy 437 ATGATCATCGCGGTTCAGGCAATCATCTGCTCATGCTTCATTCGACGATGGGCATT 496

```

Corynebacterineae; Mycobacteriaceae; Mycobacterium.

1 (sites)  
 Biglieri, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T.  
 Use of an ordered cosmid library to deduce the genomic organization  
 of Mycobacterium leprae  
 Mol. Microbiol. 7 (2), 197-206 (1993)  
 93188700  
 8446027  
 2 (bases 1 to 36985)  
 Smith, D.R., Richerich, P., Rubenfield, M., Rice, P.W., Butler, C.,  
 Lee, H.M., Kirst, S., Gundersen, K., Abendschan, K., Xu, Q., Chung, M.,  
 Deloughery, C., Aldredge, T., Maher, J., Lundstrom, R., Tulig, C.,  
 Falls, K., Imrich, J., Torrey, D., Engelstein, M., Breton, G., Madan, D.,  
 Nietupski, R., Seitz, B., Mao, J.I. et al.  
 Multiplex sequencing of 1.5 Mb of the Mycobacterium leprae genome  
 Genome Res. 7 (8), 802-819 (1997)  
 97413161  
 9267804

Original source text: Mycobacterium leprae (clone: cosmid B1529)  
 (tissue library: Lcrist 6) DNA.  
 This sequence was generated by the Genome Sequencing Center at  
 Genome Therapeutics Corporation (Collaborative Research Division),  
 100  
 Beaver St., Waltham, MA, 02154. Please contact Doug Smith  
 (smith@eric.com) for further information. The sequence represents  
 the  
 insert of a Lorient 6 cosmid clone from a mapped set of clones  
 constructed from M. leprae genomic DNA isolated from armadillo  
 liver  
 [3]. The sequence may not represent the entire cloned insert of  
 the  
 cosmid if an overlapping region was previously sequenced from  
 another  
 clone. Coding sequences larger than 60 amino acids were predicted  
 on  
 the basis of codon usage and homology information. An attempt was  
 made  
 to locate the most probable start site based on codon usage,  
 homology,  
 the presence of a Shine-Dalgarno sequence, or overlapping orf that  
 suggested translational coupling. It is possible that the actual  
 start  
 site differs from the one selected.

FEATURES  
 source  
 1..36985  
 Location/Qualifiers  
 /organism="Mycobacterium leprae"  
 /mol\_type="genomic DNA"  
 /specific\_host="Dasypus novemcinctus"  
 /db\_xref="taxon:1769"  
 /clone="cosmid B1529"  
 /tissue\_lib="Lorient 6"  
 /notes="The liver of the armadillo was used to isolate the  
 Mycobacterium leprae."

ORIGIN  
 Query Match 12.6%; Score 164; DB 1; Length 36985;  
 Best Local Similarity 52.6%; Pred. No. 3.8e-39;  
 Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;  
 Qy 257 GCGGTCGAGATCTCAAAGTCGCTGTGGGATCGGACTGGGGTCCCTGGTTCIT 316  
 Db 8360 GCGGGACGTAATTCGCCCGCGGATCGCGGTGGGCTTAAGTATGTTGTTCTCGTC 8419  
 Qy 317 TTGGGATTGCTCTAAGCCCATGGGTTGGTATCATCTCTGTCAGGTTTATGGCTGCA 376  
 Db 8420 GCACGCTGTGTTTGTCTCGGAACTCTGGTGTCTTGTGCGCTGGCCATTTTCGTT 8479  
 Qy 377 GCAACATGGAAGTTGGTAGCAGACTTAAGAGCGGCTATCATTTGCCACTGCCGATT 436  
 Db 8480 GCTAGCCATGAGTGTGCGCGGCTACGGAAGCTGGATATGTGATTCGGGTATCCCG 8539  
 Qy 437 ATGATCATCGCGGTTCAGGCAATCATCTGCTCATGCTTCATTCGACGATGGGCATT 496



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:28:22 ; Search time 53 Seconds  
(without alignments)  
1583.332 Million cell updates/sec

Title: US-09-853-641-2  
Perfect score: 1563  
Sequence: 1 MNEPEQHRSRMKPKXNA.....LPAMVTLILSVSSYPS 297

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	297	5	AAG80226 C. glutam
2	1495	95.6	285	4	AAB76559 Coryneb
3	1495	95.6	285	4	AAB80192 Coryneb
4	1495	95.6	285	4	AAG91962 C glutami
5	1064	68.1	292	6	ABU26046 Protein e
6	725.5	46.4	311	6	ABU34252 Protein e
7	719	46.0	306	6	ABU34659 Protein e
8	719	46.0	306	6	ABU36865 Protein e
9	714	45.7	312	6	ABU35960 Protein e
10	498.5	31.9	305	6	ABM65084 Propionib
11	428	27.4	228	4	AAU65884 Propionib
12	428	27.4	228	5	ABM62403 Propionib
13	350.5	22.4	262	5	ABP66169 Bifidobac
14	310	19.8	271	6	ABU41868 Protein e
15	308	19.7	285	6	ABU28067 Protein e
16	306	19.6	79	5	ABP31480 Human syn
17	303	19.4	273	6	ABU21094 Protein e
18	299	19.1	271	6	ABU39633 Protein e
19	299	19.1	282	6	ABU50039 Protein e
20	298	19.1	280	6	ABU49494 Protein e
21	298	19.1	285	6	ABU48233 Protein e
22	298	19.1	285	6	ABU47356 Protein e
23	295	18.9	263	6	ABU31563 Protein e
24	292.5	18.7	248	6	ABU45081 Protein e
25	291.5	18.7	265	3	AAU75113 Neisseria

ALIGNMENTS

RESULT 1  
AAG80226  
ID AAG80226 standard; protein; 297 AA.  
XX AC AAG80226;  
XX AC  
DT 22-JAN-2002 (first entry)  
XX AC  
DE C. glutamicum ATCC 13032 cdsA protein.  
XX AC  
KW Coryneform bacterium; cdsA; phosphatidate-cytidyltransferase; very slow  
KW L-lysine production; amino acid production; medicine; animal nutrition;  
KW growth rate.  
XX AC  
OS Corynebacterium glutamicum.  
XX AC  
FN DE10021828-Al.  
XX AC  
PD 08-NOV-2001.  
XX AC  
PF 04-MAY-2000; 2000DE-01021828.  
XX AC  
PR 04-MAY-2000; 2000DE-01021828.  
XX AC  
PA (DEGS ) DEGUSSA AG.  
XX AC (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.  
XX AC  
PI Nampoothiri M, Moeckel B, Pfeifferle W, Eggeling L, Sahm H;  
XX AC  
DR WPI; 2002-018672/03.  
XX AC  
PT N-PSDB; AAI68796.  
XX AC  
PT New mutant coryneform bacterium, useful for production of amino acids,  
XX AC especially lysine, has increased activity of phosphatidate-cytidyl  
XX AC transferase.  
XX AC  
PS Claim 10; Page 12-13; 16pp; German.  
XX AC  
CC This invention describes a novel genetically modified coryneform  
CC bacterium (A) in which activity of the cdsA gene, encoding phosphatide-  
CC cytidyl transferase, is increased (A) are used for fermentative  
CC production of amino acids, especially L-lysine, useful in human medicine,  
CC animal nutrition and pharmaceuticals. Nucleic acid, or fragments, derived  
CC from the cdsA gene are used (i) as primers for polymerase chain reaction  
CC of cdsA genes or (ii) as probes for isolating cDNA and genes that are  
CC highly homologous with cdsA. Overexpression of the cdsA gene results in  
CC increased production of amino acids, also more rapid growth to higher  
CC cell densities. This sequence represents the Corynebacterium glutamicum

full pre mature  
mature damaged  
very slow

Abp80482 N. gonorr  
Abp79682 N. gonorr  
Abu37262 Protein e  
Abu37670 Protein e  
Abu38502 Protein e  
Abu33201 Protein e  
Abu40735 Protein e  
Aay75112 Neisseria  
Abu39540 Protein e  
Abm67561 Phototrab  
Abu30427 Protein e  
Abu19701 Protein e  
Abu22263 Protein e  
Abm8605 Phototrab  
Ada35970 Acinetoba  
Abu16846 Protein e  
Abb92745 Herbicida  
Abb92123 Herbicida  
Abu43538 Protein e  
Abu26650 Protein e

```
CC ATCC 13032 phosphatidate-cytidylyl transferase (cdsA) described in the
CC method of the invention
XX
XX
SQ Sequence 297 AA;
Query Match 100.0%; Score 1563; DB 5; Length 297;
Best Local Similarity 100.0%; Pred. No. 8.3e-169;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNEPEQHRSWRMPKPKNAGRDLLKAAIATGIGLVLLGIVLSPGWYLLVAGFVAAA 60
DB 1 MNEPEQHRSWRMPKPKNAGRDLLKAAIATGIGLVLLGIVLSPGWYLLVAGFVAAA 60
QY 61 TWEVGSRLKEGSHLPLPIMIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFNGT 120
DB 61 TWEVGSRLKEGSHLPLPIMIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFNGT 120
QY 121 EKEARNYLRTSVGIFVLTWTLPLFGSFAAMLSLMNNNSIPGTYFILTFMLCVIASDVGY 180
DB 121 EKEARNYLRTSVGIFVLTWTLPLFGSFAAMLSLMNNNSIPGTYFILTFMLCVIASDVGY 180
QY 181 TAGVFFGSHPWAPLPSPKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVC 240
DB 181 TAGVFFGSHPWAPLPSPKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVC 240
QY 241 ATGLGVESQFGRDLGIKDMNSLLPGHGLMDRLDGMPLPAAMVTWLLSVISSYPS 297
DB 241 ATGLGVESQFGRDLGIKDMNSLLPGHGLMDRLDGMPLPAAMVTWLLSVISSYPS 297
RESULT 2
AAB76559
ID AAB76559 standard; protein; 285 AA.
AC AAB76559;
DT 11-APR-2001 (first entry)
XX
XX
DE Corynebacterium glutamicum MCT protein SEQ ID NO:100.
XX
XX
XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
XX membrane construction and membrane transport protein; petroleum spill;
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;
XX identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.
XX
XX
OS Corynebacterium glutamicum.
XX
XX
XX WO200100805-A2.
XX
XX
XX 04-JAN-2001.
XX
XX
XX 23-JUN-2000; 2000WO-IB000926.
XX
XX
XX 25-JUN-1999; 99US-0141031P.
XX
XX 08-JUL-1999; 99DE-01031454.
XX
XX 08-JUL-1999; 99DE-01031478.
XX
XX 08-JUL-1999; 99DE-01031563.
XX
XX 09-JUL-1999; 99DE-01032122.
XX
XX 09-JUL-1999; 99DE-01032124.
XX
XX 09-JUL-1999; 99DE-01032125.
XX
XX 09-JUL-1999; 99DE-01032128.
XX
XX 09-JUL-1999; 99DE-01032180.
XX
XX 09-JUL-1999; 99DE-01032182.
XX
XX 09-JUL-1999; 99DE-01032190.
XX
XX 09-JUL-1999; 99DE-01032191.
XX
XX 09-JUL-1999; 99DE-01032203.
XX
XX 09-JUL-1999; 99DE-01032212.
XX
XX 09-JUL-1999; 99DE-01032227.
XX
XX 09-JUL-1999; 99DE-01032228.
XX
XX 09-JUL-1999; 99DE-01032228.
XX
XX 09-JUL-1999; 99DE-01032230.
XX
XX 14-JUL-1999; 99DE-01032927.
```

```
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040830.
PR 27-AUG-1999; 99DE-01040831.
PR 27-AUG-1999; 99DE-01040832.
PR 27-AUG-1999; 99DE-01040833.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041395.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042078.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042088.
XX
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;
XX
XX WPI; 2001-071486/08.
XX N-PSDB; AAF67792.
XX
XX Corynebacterium glutamicum nucleic acids encoding membrane construction
XX and membrane transport proteins or their portions, useful for typing or
XX identifying C. glutamicum or related bacteria, and as markers for
XX transformation.
XX
XX Claim 20; Page 324; 1119pp; English.
XX
XX
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
XX construction and membrane transport (MCT) proteins given in AAB76510 to
XX AAB76847. The MCT nucleic acids and proteins are useful in the
XX identification of microorganisms which can be used to produce fine
XX chemicals, for modulating fine chemical production in C. glutamicum or
XX related bacteria (e.g. Brevibacterium lactofermentum), the typing or
XX identification of C. glutamicum or related bacteria, as reference points
XX for mapping C. glutamicum genome, and as markers for transformation.
XX AAF68082 and AAF68082 represent sequencing primers which are used in an
XX example from the present invention
XX
XX
XX SQ Sequence 285 AA;
Query Match 95.6%; Score 1495; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 4.2e-161;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 MPKPKNAGRDLLKAAIATGIGLVLLGIVLSPGWYLLVAGFVAAAATWEVGSRLKEGG 72
DB 1 MPKPKNAGRDLLKAAIATGIGLVLLGIVLSPGWYLLVAGFVAAAATWEVGSRLKEGG 60
QY 73 YHLPLPIMIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFNGTEKEARNYLRTDS 132
DB 61 YHLPLPIMIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFNGTEKEARNYLRTDS 120
QY 133 VGIFVLTWTLPLFGSFAAMLSLMNNNSIPGTYFILTFMLCVIASDVGYTAGVFFGSHPMA 192
DB 121 VGIFVLTWTLPLFGSFAAMLSLMNNNSIPGTYFILTFMLCVIASDVGYTAGVFFGSHPMA 180
QY 193 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVCATLGDVLESQF 252
DB 181 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVCATLGDVLESQF 240
QY 253 RDLGIKDMNSLLPGHGLMDRLDGMPLPAAMVTWLLSVISSYPS 297
DB 241 RDLGIKDMNSLLPGHGLMDRLDGMPLPAAMVTWLLSVISSYPS 285
RESULT 3
AAB80192
ID AAB80192 standard; protein; 285 AA.
XX
```

[illegible]

XX 18-DEC-2000; 2000EP-00127688.  
XX 16-DEC-1999; 99JP-00377484.  
XX 07-APR-2000; 2000JP-00159162.  
XX 03-AUG-2000; 2000JP-00280988.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.  
XX N-PSDB; AAH67181.  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
XX mutation point of a gene, measuring expression of a gene, analyzing  
XX expression profile or pattern of a gene and identifying homologous gene.  
XX Claim 17; SEQ ID NO 5716; 246pp + Sequence Listing; English.  
XX The present invention provides a number of nucleotide and protein  
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
XX are useful for identifying the mutation point of a gene derived from a  
XX mutant of coryneform bacterium, measuring expression amount and analyzing  
XX the expression profile or expression pattern of a gene derived from  
XX Coryneform bacterium, and identifying a homologue of a gene derived from  
XX coryneform bacterium. Coryneform bacteria are useful for producing amino  
XX acids, nucleic acids, vitamins, saccharides and organic acids,  
XX particularly L-lysine. The present sequence is a protein described in the  
XX exemplification of the invention. Note: The sequence data for this patent  
XX did not form part of the printed specification, but was obtained in  
XX electronic format directly from the European Patent Office  
XX Sequence 285 AA;  
XX Query Match 95.6%; Score 1495; DB 4; Length 285;  
XX Best Local Similarity 100.0%; Pred. No. 4.2e-161;  
XX Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX 13 MPKPKNNAGRLKAAIAVGIGLVLLGVLSPGWVILVAGFMAATWEVGRLEGG 72  
XX 1 MPKPKNNAGRLKAAIAVGIGLVLLGVLSPGWVILVAGFMAATWEVGRLEGG 60  
XX 73 YHLPPLMIGGQAIILSWPFGTMGILAGFVATVLMVFRIFYNGTEKEARNYLRDTS 132  
XX 61 YHLPPLMIGGQAIILSWPFGTMGILAGFVATVLMVFRIFYNGTEKEARNYLRDTS 120  
XX 133 VGIFVLTWIPFGSFAAMLSLMONNSTPGTYFILTFLMVCVADVGVIAGVFGSHPM 192  
XX 121 VGIFVLTWIPFGSFAAMLSLMONNSTPGTYFILTFLMVCVADVGVIAGVFGSHPM 180  
XX 193 PLVSPKKSWEAGFAGSIVLGSVTGALSVEHLLDHHWMMGVILGCVLVVCAITGLDVSQPK 252  
XX 181 PLVSPKKSWEAGFAGSIVLGSVTGALSVEHLLDHHWMMGVILGCVLVVCAITGLDVSQPK 240  
XX 253 RDLGKIDMSNLLPHGGIMRDLGMLPAAVMVTLILSVSSSYPS 297  
XX 241 RDLGKIDMSNLLPHGGIMRDLGMLPAAVMVTLILSVSSSYPS 285  
XX RESULT 5  
XX ABU26046  
XX ID ABU26046 standard; protein; 292 AA.  
XX AC ABU26046;  
XX 19-JUN-2003 (first entry)  
XX Protein encoded by Prokaryotic essential gene #11573.  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Corynebacterium diphtheriae.  
XX PN WO200277183-A2.  
XX XX 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362599P.  
XX XX (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
XX DR N-PSDB; ACA29916.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 53970; 1766pp; English.  
XX XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. the present sequence is encoded by one of  
XX the target prokaryotic essential genes. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 292 AA;  
XX Query Match 68.1%; Score 1064; DB 6; Length 292;  
XX Best Local Similarity 70.5%; Pred. No. 4.7e-112;  
XX Matches 198; Conservative 33; Mismatches 50; Indels 0; Gaps 0;  
XX 13 MPKPKNNAGRLKAAIAVGIGLVLLGVLSPGWVILVAGFMAATWEVGRLEGG 72  
XX 11 LPKPKNSAGRLKAAISVGIGLGVLLAIFVIFPGWVPLVIAIAVATWEVERLEAG 70  
XX 73 YHLPPLMIGGQAIILSWPFGTMGILAGFVATVLMVFRIFYNGTEKEARNYLRDTS 132  
XX 71 YLLQRWMLIGGQVLMWLSNPPGPKGLVAGFVGVLATMFGRLFHGGRSPPKNYLRDTA 130

QY 133 VGIFVLTWIPFGSFAWLSLMQNNISIPGTYRILTFMCLVIAADVGVIAGVFFGSHPVA 192  
Db 131 VAIFVLTWIPFGSFAWLSLTFETAPGKFIIVTFMCLVIAADVGVIAGVFFGSHPVA 190  
QY 193 PLVSPKSWGFGAGSIVLGSVTGALSIVHFLDHHWMMGVILGALVVCATLGLDLVESQPK 252  
Db 191 PAVSPKSWGFGSIVGVFGMIVGAFVSYLLGHQWMMGLVGLFGLVCATLGLDLVESQPK 250  
QY 253 RDLGIKMSNLLPHGGLMDRLDGMVIPAAMVTWILSVISS 293  
Db 251 RELGIKMSAILPHGGLMDRLDGMFSAAMVTWILSVISS 291

RESULT 6  
ABU34252  
ID ABU34252 standard; protein; 311 AA.  
XX AC ABU34252;  
DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #19779.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Mycobacterium avium.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
XX DR N-PSDB; ACA38122.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids, required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 62176; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 311 AA;

Query Match 46.4%; Score 725.5; DB 6; Length 311;  
Best Local Similarity 47.5%; Pred. No. 1.6e-73;  
Matches 145; Conservative 47; Mismatches 94; Indels 19; Gaps 4;  
QY 2 NEPE-----QHRSRMKPKNNAGRDLLAAIAVIGLGVILGLVSPWGYILVA 54  
Db 10 DEPEHAVENTTEGAAGQRAKKTTSRAGRDLLAAIAVIGLGVILGLVSPWGYILVA 69  
QY 55 GFMAATWVGSRLEGGYHLPLPIMILGGQALILSWPFGTMGILAGFVATVLMVPR 114  
Db 70 MAILVASHVEVRLEAGYIVIPILAGGQUTVMTWTPFHAAGALAGGVTVACLFWR 129  
QY 115 IFYNGTEKE-----ARNYLRTDSVGIFVLTWLPFGSFAAMLSLMQNNISIPGYFL 166  
Db 130 LFMQDNKRPPEPPAGSPSANYLRDASATVFLACWVLEFASFAALLVYPADGA--GRVFCL 187  
QY 167 TFMCLVIAADVGVIAGVFFGSHPMAPLVSPKSWGFGAGSIVLGSVTCALSVHFLDDH 226  
Db 188 --MITVASDVGGYAVGVNLFPGHWPVPAISPSPKSWGFGAGSIVLGSVTCALSVHFLDDH 245  
QY 227 WMMGVILGALVVCATLGLDLVESQPKRDLGIDKMSNLLPHGGLMDRLDGMFSAAMVTWL 286  
Db 246 PWGALLGVVLTCTLGLDLVESQPKRDLGIDKMSNLLPHGGLMDRLDGMFSAAMVTWL 305  
QY 287 ILSVI 291  
Db 306 VLTIV 310

RESULT 7  
ABU34659  
ID ABU34659 standard; protein; 306 AA.  
XX AC ABU34659;  
XX AC ABU34659;  
DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #20186.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Mycobacterium bovis.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

```

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
DR N-PSDB; ACA38529.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 62583; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 46.0%; Score 719; DB 6; Length 306;
XX Best Local Similarity 47.7%; Pred. NO. 8.8e-73;
XX Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;
XX
QY 2 NEPEQHHSMEM--PKPQNNAGRDLCRAIAVIGLGVLLGVLSPWGVILVAGFMAA 59
Db 10 NPAQPPAKGAQQPQATETSRAGDRRAIVVGLSIGLVIAVLVFFVRVVAIVATVL 69
QY 60 ATWEVGSRLKEGGYHLPIMIGQQAIIILSWPFGTWGILASPVATVLVLMVFRIPY-- 117
Db 70 ATHEVRELREAGYLLPVIPLIGQAQVWLTFPGAVGALAGFGMVVCMWLFLPMQD 129
QY 118 -----NCTEKARNYLDTESVGLVTLPLFGSFAAMLNMONNSIPQTYILTFMLC 171
Db 130 SVTRPTTGAPSPGNYLSDVSATVFLAVVPLFCFGAMLVYPEN----GSGWVFCNMIA 185
QY 172 VIASDVGGYIAGVFGGHPMAPLSPKKSNEGFPAGSIVLGSVTVGALSVEHLDDHWMGV 231
Db 186 VIASDVGGYAGVLFGRHPVPTISPKKSNEGFPAGSLVCGITATITATLVGKTPWGA 245
QY 232 ILGALVVCATLGLDVRSGQKRDRLGIDKMNLLPFGHGLMDRLDGLMPAAWTLVLSVI 291
Db 246 LLGLVFLVLTALGLDVRSGQKRDRLGIDKMNLLPFGHGLMDRLDGLMPAAWTLVLTLL 305

RESULT 8
ABU36865
ID ABU36865 standard; protein; 306 AA.

```

```

XX AC ABU36865;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #22392.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Mycobacterium tuberculosis.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX PI WPI: 2003-029926/02.
XX DR N-PSDB; ACA40735.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 64789; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 46.0%; Score 719; DB 6; Length 306;
XX Best Local Similarity 47.7%; Pred. NO. 8.8e-73;

```







Db 112 DTGGYAVGVLCGKHKLAPRISPKKSWEGFAGSVITAAAFVGVACWACLGGLLSAPWAGIVLG 171

Qy 236 ALVVCATLGLDVLVESQFKRDGLGKIDMSNLLPGHGLMDRLDGLMPLPAAWTVLILSVI 291

Db 172 VLALTGTAGDLVESMIKRDAGIKDMSNLFPGHGVMDRLDVLFSAPFAMVMSLV 227

RESULT 12

ABM62403

ID ABM62403 standard; protein; 228 AA.

XX AC

XX ABM62403;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes membrane-related polypeptide #27079.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX OS

XX WO2003033515-A1.

XX PN

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX PF

XX 15-OCT-2001; 2001US-00978825.

XX PR

XX (CORI-) CORIXA CORP.

XX PA

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX PI Barth B, Vallieve-bouglass J;

XX WPI; 2003-381789/36.

XX DR N-PSDB; ACF64627.

XX PT

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

XX or for stimulating an immune response specific for a P. acnes protein.

XX PS

XX Claim 3d; SEQ ID NO 27079; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

XX encoding a Propionibacterium acnes protein. The invention also relates to

XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

XX immunogenic fragments of P. acnes polypeptides. The invention

XX additionally encompasses expression vectors and host cells comprising a

XX polynucleotide of the invention; antibodies against polypeptides of the

XX invention; fusion proteins comprising a polypeptide of the invention; a

XX method for stimulating an immune response specific for a P. acnes

XX polypeptide and an isolated T cell population comprising T cells prepared

XX via this method; a vaccine composition (comprising P. acnes polypeptides,

XX polynucleotides, antibodies, fusion proteins, T cell populations, or

XX antigen-presenting cells that express the polypeptide); a method and kit

XX for detecting or determining the presence or absence of P. acnes in a

XX patient; and a method for inhibiting the development of P. acnes in a

XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

XX proteins, T cell populations or antigen-presenting cells that express the

XX polypeptides are useful for diagnosing, preventing or treating acne

XX vulgaris, or for stimulating an immune response specific for a P. acnes

XX protein. The polynucleotides can also be used as probes or primers for

XX nucleic acid hybridization. The vaccine composition is useful for the

XX stimulation of an immune response against P. acnes, or for treating acne,

XX and the kit is useful for performing a diagnostic assay. The present

XX sequence represents a specifically claimed polypeptide which is predicted

XX to be encoded by an ORF (open reading frame) contained within the P.

XX acnes polynucleotides of the invention. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 228 AA;

SQ

Query Match 27.4%; Score 428; DB 6; Length 228;

Best Local Similarity 46.6%; Pred. No. 7.4e-40;

Matches 82; Conservative 32; Mismatches 56; Indels 6; Gaps 3;

QY 118 NG--TEKEARNYLRDTSVGI FVL TWIPLEFGSPAAMLSLMONNSIPGTVPILTFMLCVIAS 175

Db 56 NGVATARGSHGFVKDAASLFTIAYLPLLGCP--VPLMMGDD-CGSRRIATWILSVAS 111

QY 176 DVGGYIAGVFFGSHPMAPLVSPKKSWEFGFAGSIVLGSVTVGALSVMHFLDHHVMGVILGC 235

Db 112 DTGGYAVGVLCGKHKLAPRISPKKSWEGFAGSVITAAAFVGVACWACLGGLLSAPWAGIVLG 171

QY 236 ALVVCATLGLDVLVESQFKRDGLGKIDMSNLLPGHGLMDRLDGLMPLPAAWTVLILSVI 291

Db 172 VLALTGTAGDLVESMIKRDAGIKDMSNLFPGHGVMDRLDVLFSAPFAMVMSLV 227

RESULT 13

ABP66169

ID ABP66169 standard; protein; 362 AA.

XX AC

XX ABP66169;

XX 19-NOV-2002 (first entry)

XX DT

XX Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:913.

XX DE

XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;

XX KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;

XX KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;

XX KW rotavirus; food composition; pharmaceutical composition.

XX OS

XX Bifidobacterium longum.

XX PN

XX EP1227152-A1.

XX 31-JUL-2002.

XX PD

XX 30-JAN-2001; 2001EP-00102050.

XX PF

XX 30-JAN-2001; 2001EP-00102050.

XX PR

XX (NEST ) SOC PROD NESTLE SA.

XX PA

XX WPI; 2002-668397/72.

XX DR

XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as

XX a probe or primer for detecting and/or identifying Bifidobacterium longum

XX in a biological sample.

XX PT

XX Claim 3; SEQ ID NO 913; 80pp; English.

XX PS

XX The present invention describes a polynucleotide (I) comprising a

XX sequence of a Bifidobacterium genome selected from the nucleotide

XX sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at

XX least 90% identity or which hybridises with the sequences given in

XX ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a

XX fusion protein, comprising a sequence selected from 1097 sequences given

XX in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a

XX heterologous polypeptide. (I) has antidiarrheic and antibacterial

XX activities, and can be used as an inhibitor of salmonella. (I) (which is

XX a probe) is useful for the detection and/or identification of

XX Bifidobacterium longum in a biological sample. A carrier containing the

XX lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be

XX used for preventing and/or treating diarrhoea brought about by pathogenic

XX bacteria and/or rotavirus. The carrier is a food composition selected

XX from milk, yogurt, curd, cheese, fermented milks, milk based fermented

XX products, ice-creams, fermented cereal based products, milk based

XX powders, infant formula, pet food or a pharmaceutical composition

XX selected from tablets, liquid bacterial suspensions, dried oral

supplement, wet oral supplement, dry tube feeding or wet tube feeding.  
(i) is useful in DNA arrays or chips to carry out analysis of the expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent Bifidobacterium related nucleotide sequences given in the Sequence Listing from the present invention but not mentioned further within the specification. N.B. The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 16:56:56 ; Search time 861 Seconds  
(without alignments)  
7408.351 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300  
Sequence: 1 gaactcgttctgcgaagga.....aggaaaccgattccgctgg 1300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	1300	12	US-09-853-641-1
2	1300	100.0	3309400	9	US-09-738-626-1
3	978	75.2	978	13	US-10-627-476-99
4	855	65.8	855	9	US-09-738-626-2216
5	342.2	26.3	876	13	US-10-282-122A-17786
6	191.8	14.8	918	13	US-10-282-122A-26399
7	191.8	14.8	921	13	US-10-282-122A-28605
8	168.8	13.0	933	13	US-10-282-122A-25392
9	164	12.6	939	13	US-10-282-122A-27700
10	123.4	9.5	297	11	US-09-864-408A-905
11	117	9.0	258	9	US-09-738-626-2215
12	110.2	8.5	1176	15	US-10-156-761-2610
13	110.2	8.5	9025608	15	US-10-156-761-1
14	69.8	5.4	2256645	17	US-10-470-565-1

15	68.4	5.3	813	13	US-10-282-122A-33608
16	67.2	5.2	813	13	US-10-282-122A-31373
17	60.8	4.7	2731748	17	US-10-297-465A-1
18	59	4.5	843	13	US-10-282-122A-41234
19	58.8	4.5	936	9	US-09-815-243-7776
20	58.8	4.5	936	13	US-10-282-122A-30203
21	58.4	4.5	2017	16	US-10-381-779-95
22	58.2	4.5	789	13	US-10-282-122A-23303
23	58.2	4.5	795	13	US-10-282-122A-24941
24	57.2	4.4	879	16	US-10-369-493-39691
25	57.2	4.4	879	16	US-10-369-493-40043
26	57.2	4.4	927	16	US-10-369-493-39316
27	56.6	4.4	819	13	US-10-282-122A-11441
28	55.8	4.3	855	13	US-10-282-122A-15294
29	55.2	4.2	789	13	US-10-282-122A-24244
30	55.2	4.2	1733	16	US-10-398-221-3425
31	55	4.2	819	13	US-10-282-122A-14003
32	54.6	4.2	748	13	US-10-282-122A-36821
33	54.6	4.2	858	13	US-10-282-122A-39096
34	54.4	4.2	7060	13	US-10-194-163-479
35	53.2	4.1	846	13	US-10-282-122A-26860
36	53.2	4.1	99629	13	US-10-672-787-37
37	53	4.1	858	13	US-10-282-122A-39973
38	52	4.0	792	13	US-10-282-122A-36212
39	51	3.9	795	13	US-10-282-122A-29002
40	50.8	3.9	861	13	US-10-282-122A-32475
41	50	3.8	780	13	US-10-282-122A-12345
42	49.6	3.8	855	13	US-10-282-122A-19807
43	49	3.8	849	13	US-10-282-122A-41779
44	47.8	3.7	798	13	US-10-282-122A-29410
45	47.4	3.6	795	13	US-10-282-122A-38585

ALIGNMENTS

RESULT 1

US-09-853-641-1  
; Sequence 1, Application US/09853641  
; Publication No. US20040092710A1  
; GENERAL INFORMATION:  
; APPLICANT: NAMPOOTHIRI, Madhavan  
; TITLE OF INVENTION: Nucleotide Sequences Coding for the CdsA Gene  
; FILE REFERENCE: 032301 WD 1171  
; CURRENT APPLICATION NUMBER: US/09/853,641  
; CURRENT FILING DATE: 2001-08-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1300  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (200)..(1090)  
; OTHER INFORMATION: cdsa gene  
; NAME/KEY: RBS  
; LOCATION: (187)..(195)  
; OTHER INFORMATION:  
US-09-853-641-1

Query Match	100.0%	Score 1300;	DB 12;	Length 1300;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GAAGTCGTTCTCGCAAGGAAAGGAACGTATGATGAGGTCTAGAGACCTTTATCGCAATG	60	
Db	1	GAAGTCGTTCTCGCAAGGAAAGGAACGTATGAGGTCTAGAGACCTTTATCGCAATG	60	
Qy	61	GCTTAAGAATACGTCGTTTCAACGTCGATGGCGGGGAAACGACGCTTTCTTTGCTT	120	
Db	61	GCTTAAGAATACGTCGTTTCAACGTCGATGGCGGGGAAACGACGCTTTCTTTGCTT	120	

121 GCAAGAGTGTGTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTGCATTTGTTA 180  
121 GCAAGAGTGTGTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTGCATTTGTTA 180  
181 GAAACTTCAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 240  
181 GAAACTTCAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 240  
241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300  
241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300  
301 GGGGTCTCTGTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360  
301 GGGGTCTCTGTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360  
361 AGTTTATGCTGCTGAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCA 420  
361 AGTTTATGCTGCTGAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCA 420  
421 TTTGCCACTGCCATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCGCAT 480  
421 TTTGCCACTGCCATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCGCAT 480  
481 TGGCAGATGGCATTTGGCGCTTTTGGCCACTGTGTGGTGGTGTGATTTCCG 540  
481 TGGCAGATGGCATTTGGCGCTTTTGGCCACTGTGTGGTGGTGTGATTTCCG 540  
541 AATTTTCTACATGCGCAAGAAAAAGCCCGCACTATTAGGGACACCTCTGTGGG 600  
541 AATTTTCTACATGCGCAAGAAAAAGCCCGCACTATTAGGGACACCTCTGTGGG 600  
601 CATCTCTGCTCACTGATTCATTTGTTGGAGCTTCCTGGCATGCTGTGCTGAT 660  
601 CATCTCTGCTCACTGATTCATTTGTTGGAGCTTCCTGGCATGCTGTGCTGAT 660  
661 GCAAAACAATTCATCCCGGGTACATATTTTCACTTGAAGCTTCATGCTGTGTGATCGC 720  
661 GCAAAACAATTCATCCCGGGTACATATTTTCACTTGAAGCTTCATGCTGTGTGATCGC 720  
721 ATCGATGTGGCGGGTATATCGCGGTGTGTTTGGATCGCAACCAATGGCGCGTT 780  
721 ATCGATGTGGCGGGTATATCGCGGTGTGTTTGGATCGCAACCAATGGCGCGTT 780  
781 GGTGAGTCGGAAGAGTCTTGGGAAGGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 840  
781 GGTGAGTCGGAAGAGTCTTGGGAAGGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 840  
841 TGGTGCATCTAGTGTCACTTCTCTGCTGATCACCATGTTGGTGGATGGTGTGATTTGGG 900  
841 TGGTGCATCTAGTGTCACTTCTCTGCTGATCACCATGTTGGTGGATGGTGTGATTTGGG 900  
901 TTGTCCTAGTGTGTCGCGCACGTTGGTGGATCTTGGTGGATGGTGTGATTTGGG 960  
901 TTGTCCTAGTGTGTCGCGCACGTTGGTGGATCTTGGTGGATGGTGTGATTTGGG 960  
961 TTTGGGCTCAAGGATATGTCGAACTTTCTCCAGGCCACCGCGGATGATGGACCGTTT 1020  
961 TTTGGGCTCAAGGATATGTCGAACTTTCTCCAGGCCACCGCGGATGATGGACCGTTT 1020  
1021 GGATGGCATCTCCCGCGCGGATGGTGAAGTGGTGTGATCTGAGTGTGATCAGCAGTTC 1080  
1021 GGATGGCATCTCCCGCGCGGATGGTGAAGTGGTGTGATCTGAGTGTGATCAGCAGTTC 1080  
1081 GTATCCGTCGTAAAGCTTTGGGCGACCTTTAAGTTCAAAAACTTTGAAGGCGCTGAGTG 1140  
1081 GTATCCGTCGTAAAGCTTTGGGCGACCTTTAAGTTCAAAAACTTTGAAGGCGCTGAGTG 1140  
1141 CATACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 1200  
1141 CATACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 1200  
1201 TTCTTTAACTGGCGAGGTACTCAACATGCGCAGCCCAACCAAGCGCCATATCAATG 1260

1201 TTCTTAACCTGGCGAGGTACTCAAAACATGCGCAGCCCAACAGCGCCATTAATCAATG 1260  
1261 CACCGGTAAATGGCTGTAGTAGGAACCGATTCCGGCTGG 1300  
1261 CACCGGTAAATGGCTGTAGTAGGAACCGATTCCGGCTGG 1300

RESULT 2  
US-09-738-626-1/c  
; Sequence 1, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent in ver. 3.0  
; SEQ ID NO 1  
; LENGTH: 3309400  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-1

Query Match 100.0%; Score 1300; DB 9; Length 3309400;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTTGCTCGCAAGGAAAGGAACCTGATGGAGTCTAGAAGACCTTTATCGCAATG 60  
DB 2137324 GAAGTCGTTGCTCGCAAGGAAAGGAACCTGATGGAGTCTAGAAGACCTTTATCGCAATG 2137265  
QY 61 GCTTAAGAAATACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTTGGCTT 120  
DB 2137264 GCTTAAGAAATACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTTGGCTT 2137205  
QY 121 GCAAGAGTCTTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTGCATTTGTTA 180  
DB 2137204 GCAAGAGTCTTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTGCATTTGTTA 2137145  
QY 181 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 240  
DB 2137144 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 2137085  
QY 241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300  
DB 2137084 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 2137025  
QY 301 GGGGTCTCTGTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360  
DB 2137024 GGGGTCTCTGTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 2136965  
QY 361 AGTTTATGCTGCTGAGCAACATGGAAGTTGGTAGCAGCTTAAAGAGCGCGCTATCA 420  
DB 2136964 AGTTTATGCTGCTGAGCAACATGGAAGTTGGTAGCAGCTTAAAGAGCGCGCTATCA 2136905



```
QY 421 TTTGCCAGTCCGATTATGATCATCGCGGTGTCAGGCAATCATCTGGTGTGTCAGGCAATT 480
Db 2136904 TTTGCCAGTCCGATTATGATCATCGCGGTGTCAGGCAATCATCTGGTGTGTCAGGCAATT 2136845
QY 481 TGGCAGATGGCAATTTGGCGTCTTTTGTGGGCACTGTTGGTGTGTAATGATTTCCG 540
Db 2136944 TGGCAGATGGCAATTTGGCGTCTTTTGTGGGCACTGTTGGTGTGTAATGATTTCCG 2136785
QY 541 AATTTTCTAATGGCAGGAAAGAAAGCCGCAACTATTTGAGGACACCTCTGTGG 600
Db 2136784 AATTTTCTAATGGCAGGAAAGAAAGCCGCAACTATTTGAGGACACCTCTGTGG 2136725
QY 601 CATCTTGGTGTGTCACCTGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGTGCTGAT 660
Db 2136724 CATCTTGGTGTGTCACCTGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGTGCTGAT 2136665
QY 661 GCAAAACAATTCATCCCGGGTACATATTTTGAAGCTTCATGCTGTGTCGTCGC 720
Db 2136664 GCAAAACAATTCATCCCGGGTACATATTTTGAAGCTTCATGCTGTGTCGTCGC 2136605
QY 721 ATCGGATGTGGCGGTATATCGCGGTGTCTTTTGGATCGCACCATGCGCGCGTT 780
Db 2136604 ATCGGATGTGGCGGTATATCGCGGTGTCTTTTGGATCGCACCATGCGCGCGTT 2136545
QY 781 GGTGAGTCCGAAAGTCTTGGGAAGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 840
Db 2136544 GGTGAGTCCGAAAGTCTTGGGAAGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 2136485
QY 841 TGGTGCACTAGTGTCTCACTTCCTGCTCGATCACACTGGTGGATGCTGTGATCTGG 900
Db 2136484 TGGTGCACTAGTGTCTCACTTCCTGCTCGATCACACTGGTGGATGCTGTGATCTGG 2136425
QY 901 TTGTGCCCTAGTGTGTGCGCCACGTTGGGTGACTTTGTTGAGTGGCAGTTCAAACGCGA 960
Db 2136424 TTGTGCCCTAGTGTGTGCGCCACGTTGGGTGACTTTGTTGAGTGGCAGTTCAAACGCGA 2136365
QY 961 TTTGGGATCAAGATATGTGGAACCTTCTTCCAGGCCACCGCGGATGATGGACGCTTT 1020
Db 2136364 TTTGGGATCAAGATATGTGGAACCTTCTTCCAGGCCACCGCGGATGATGGACGCTTT 2136305
QY 1021 GGATGGCATGCTCCGCGCGGATGATGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 1080
Db 2136304 GGATGGCATGCTCCGCGCGGATGATGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 2136245
QY 1081 GTATCCGTCGTAAGCTTGGCCAGCTTTAAGTTCAAAAACCTTGAAGGCGCTGAGGTG 1140
Db 2136244 GTATCCGTCGTAAGCTTGGCCAGCTTTAAGTTCAAAAACCTTGAAGGCGCTGAGGTG 2136185
QY 1141 CATACGTCGTCACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTACGGATTT 1200
Db 2136184 CATACGTCGTCACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTACGGATTT 2136125
QY 1201 TTTTAACTGCGCGGCTACTCAAAACATGCGCAGCCCAACCAAGCGCCCAATATCAATG 1260
Db 2136124 TTTTAACTGCGCGGCTACTCAAAACATGCGCAGCCCAACCAAGCGCCCAATATCAATG 2136065
QY 1261 CACCGGTAATGGCTGCTAGTAGGAACCGATTCGGCTGG 1300
Db 2136064 CACCGGTAATGGCTGCTAGTAGGAACCGATTCGGCTGG 2136025
```

## RESULT 3

US-10-627-476-99  
; Sequence 99, Application US/10627476  
; Publication No. US20040030116A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Mark  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schoder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Habernauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE

; TITLE OF INVENTION: TRANSPORT  
; FILE REFERENCE: BGI-125PCPN  
; CURRENT APPLICATION NUMBER: US/10/627,476  
; CURRENT FILING DATE: 2003-07-25  
; PRIOR APPLICATION NUMBER: 09/602,787  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: USSN 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931454.3  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931478.0  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931563.9  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19932122.1  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932124.8  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932125.6  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932128.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932180.9  
; PRIOR FILING DATE: 1999-07-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 678  
; SEQ ID NO 99  
; LENGTH: 978  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)...(955)  
; OTHER INFORMATION: RXA01894  
US-10-627-476-99

Query Match 75.2%; Score 978; DB 13; Length 978;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 136 AGAATTTTTCGAAAATGTCGACCATCAACAGTGCACATTTGTAGAACTTTCAAGGAGA 195
Db 1 AGAATTTTTCGAAAATGTCGACCATCAACAGTGCACATTTGTAGAACTTTCAAGGAGA 60
QY 196 ACCATGATGACCGGACCAATCACCGTCCATGAGATGCCCAACCCCAAAATAA 255
Db 61 ACCATGATGACCGGACCAATCACCGTCCATGAGATGCCCAACCCCAAAATAA 120
QY 256 TCGGGTTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATCGGACTGGGGTCTTGGTTCT 315
Db 121 TCGGGTTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATCGGACTGGGGTCTTGGTTCT 180
QY 316 TTTGGGGATTCCTAAGCCCATGGGGTTGGTACATCTCTGTGAGGTTTATGGCTGC 375
Db 181 TTTGGGGATTCCTAAGCCCATGGGGTTGGTACATCTCTGTGAGGTTTATGGCTGC 240
QY 376 AGCAACATGGGAATTTGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 435
Db 241 AGCAACATGGGAATTTGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 300
QY 436 TATGATCATCGCGGTGAGCAATCATCTGGGTGTGATGCCCATTTGGCAGCATGGGCAT 495
Db 301 TATGATCATCGCGGTGAGCAATCATCTGGGTGTGATGCCCATTTGGCAGCATGGGCAT 360
QY 496 TTTGGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
Db 361 TTTGGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 556 CACGAAAAAGAACCCCGCAACTATTGAGGGACACCTCTGTGGGATCTTGTGTGTGTGTGT 615
Db 421 CACGAAAAAGAACCCCGCAACTATTGAGGGACACCTCTGTGGGATCTTGTGTGTGTGTGT 480
QY 616 CTGGAATTCATTTGTCGGAAGCTTCGCTGCGATGCTGTGCTGATGCAAAACAATTCAT 675
```

Db 481 CTGGATTCATTTGTCGGAAGCTTCTGCTCGATGCTGCTGATGCAAAACATTCAT 540  
 QY 676 CCGGGTACATATTCATTTGACGTTGANGCTGTGTGTGATGCGATCGGATGCGGG 735  
 Db 541 CCGGGTACATATTCATTTGACGTTGATGCTGTGTGTGATGCGATCGGATGCGGG 600  
 QY 736 GTATATCGCGGGTGTGTTCTTTGGATCGACCAATGCGCGCTGTGTGAGTCCGAAGAA 795  
 Db 601 GTATATCGCGGGTGTGTTCTTTGGATCGACCAATGCGCGCTGTGTGAGTCCGAAGAA 660  
 QY 796 GTCTTTGGGAAGGCTTTGCGGGCTCCATTGCTTTAGGATCGGTCACCTGCTGCACTCAGTGT 855  
 Db 661 GTCTTTGGGAAGGCTTTGCGGGCTCCATTGCTTTAGGATCGGTCACCTGCTGCACTCAGTGT 720  
 QY 856 TCACCTTCCTGCTCGATCACCACCTGCTGGATGCGGTGTGATCTTGGTTGCTGCTAGTTGT 915  
 Db 721 TCACCTTCCTGCTCGATCACCACCTGCTGGATGCGGTGTGATCTTGGTTGCTGCTAGTTGT 780  
 QY 916 GTGCGCCACGTTGGGTGACTTGGTTGAGTCGCGAGTTCAAACCGCAATTTGGGCATCAAGGA 975  
 Db 781 GTGCGCCACGTTGGGTGACTTGGTTGAGTCGCGAGTTCAAACCGCAATTTGGGCATCAAGGA 840  
 QY 976 TATGTCGAACCTTCTTCCAGCCACGCGGAGTGTGAGACCGTTTGGATGCGATGCTCCC 1035  
 Db 841 TATGTCGAACCTTCTTCCAGCCACGCGGAGTGTGAGACCGTTTGGATGCGATGCTCCC 900  
 QY 1036 GCGCGCGATGGTGAGTGTGATCTGATCTGAGTGTGATCAGCAGCTGATCTCGTGAAG 1095  
 Db 901 GCGCGCGATGGTGAGTGTGATCTGATCTGAGTGTGATCAGCAGCTGATCTCGTGAAG 960  
 QY 1096 CTTGGGCCAGCTTTAAGT 1113  
 Db 961 CTTGGGCCAGCTTTAAGT 978

## RESULT 4

US-09-738-626-2216  
 ; Sequence 2216, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 2216  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-2216

Query Match 65.8%; Score 855; DB 9; Length 855;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-275;  
 Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 ATGCCCAACCCAAAATAATATGCGGTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATC 295  
 Db 1 ATGCCCAACCCAAAATAATATGCGGTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATC 60  
 QY 296 GGAAGTGGGGTCTGCTGTTCTTTGGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTC 355  
 Db 61 GGAAGTGGGGTCTGCTGTTCTTTGGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTC 120  
 QY 356 GTTGCAAGTTTATGCTGTCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGGC 415  
 Db 121 GTTGCAAGTTTATGCTGTCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGGC 180  
 QY 416 TATCATTTGCCACTCGCATTTATGATCATCGGCGGTCAGGCAATCATCTGGCTGTGATG 475  
 Db 181 TATCATTTGCCACTCGCATTTATGATCATCGGCGGTCAGGCAATCATCTGGCTGTGATG 240  
 QY 476 CCATTTGGCAACGATGGGCATTTTGGCGTCTTTTGTGGCCACTGTGTGTGTGTGATG 535  
 Db 241 CCATTTGGCAACGATGGGCATTTTGGCGTCTTTTGTGGCCACTGTGTGTGTGTGATG 300  
 QY 536 TTCGGAATTTCTACAAATGCGCAAGAAAGAGCCCGCAACTATTTAGAGGACACCTCT 595  
 Db 301 TTCGGAATTTCTACAAATGCGCAAGAAAGAGCCCGCAACTATTTAGAGGACACCTCT 360  
 QY 596 GTGGGCATCTTCTGTGCTCACCTGGATTCCATTGTTTCGGAAGCTTCGCTCGGATGCTG 655  
 Db 361 GTGGGCATCTTCTGTGCTCACCTGGATTCCATTGTTTCGGAAGCTTCGCTCGGATGCTG 420  
 QY 656 CTGATCAAAACAAATTCATCCGGGTACATATTTCAATTTGACGTTCACTGTGTGTG 715  
 Db 421 CTGATCAAAACAAATTCATCCGGGTACATATTTCAATTTGACGTTCACTGTGTGTG 480  
 QY 716 ATCGCATCGGATGTCGGCGGTATATCGGGGTGTGTTCTTTGGATCGCACCAATGGCG 775  
 Db 481 ATCGCATCGGATGTCGGCGGTATATCGGGGTGTGTTCTTTGGATCGCACCAATGGCG 540  
 QY 776 CCGTTGGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGCTCCATTGCTTTAGGATCG 835  
 Db 541 CCGTTGGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGCTCCATTGCTTTAGGATCG 600  
 QY 836 GTCACTGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCACCTGCTGATGCTGTGATC 895  
 Db 601 GTCACTGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCACCTGCTGATGCTGTGATC 660  
 QY 896 TTGGTTGTGCCCTAGTTGTGTGCGCCAGCTTGGGTGACTTGGTTGAGTCGAGTTCAA 955  
 Db 661 TTGGTTGTGCCCTAGTTGTGTGCGCCAGCTTGGGTGACTTGGTTGAGTCGAGTTCAA 720  
 QY 956 CCGGATTTGGGCATCAAGGATATGCGAACCTTCTTCAGGCGCACGCGGATGATGAGAC 1015  
 Db 721 CCGGATTTGGGCATCAAGGATATGCGAACCTTCTTCAGGCGCACGCGGATGATGAGAC 780  
 QY 1016 CGTTTGGATGGATGCTCCGCGCCGCGATGCTGAGTGTGATCTCTGAGTGTGATCAGC 1075  
 Db 781 CGTTTGGATGGATGCTCCGCGCCGCGATGCTGAGTGTGATCTCTGAGTGTGATCAGC 840  
 QY 1076 AGCTGTATCCGTCG 1090  
 Db 841 AGCTGTATCCGTCG 855

## RESULT 5

US-10-282-122A-17786  
 ; Sequence 17786, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Walli, Daniel

```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17786
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17786

Query Match      26.3%; Score 342.2; DB 13; Length 876;
Best Local Similarity 63.1%; Pred. No. 4e-103;
Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

QY 237 TGCCCAACCCAAATAATCGGGTCAGATCTCAAGCTGCCATGCTGTGGGGATCG 296
DB 32 TGCCCTAAACCGAAGAATTCGGCTGGTCGAACCTGAAAGCCGCGATCAGCCTTGGAAATG 91

QY 297 GACTGGGGGCTCTGCTTTCTTTGGGGATGTCCTAAAGCCCATGGGGTTGGTACATCCTCG 356
DB 92 GCTTGGGGCCCTCGTATGCTCGTATCTTTGTTATTCCTTTTGGTTGGTATCCATTAG 151

QY 357 TTGCAGGTTTATGGCTCAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCT 416
DB 152 TGGCTATCGCAATCGCAGTGGCACTTGGGAGGTAGAGCGCGTTTAAATAGAAGCCGGTT 211

QY 417 ATCATTTCCCACTGCGATTATGATCATCGCGGTTCAGGCAATCATCTGGCTGTCATGGC 476
DB 212 ATCTTTTCAACGTGGTGTATGCTATATGGGACAAAGTCAATGTTGGCTTAGTTGGC 271

QY 477 CATTTGGCAGCATGGCAATTTTGGCGCTTTTGGGCCACTGTGTGGTGTCTGATGATT 536
DB 272 CATTTGGTCCGAAGGATTAGTGGCTGTTTCGTGGGGGTTGTACTCGTACGATGTTG 331

QY 537 TCCGAATTTTCTACATGGCAGGAAAGAACCCGCAACTATTAGGAGCAACCTCTG 596
DB 332 GTCGCGCTTTTCCATCATGGTCTTATGTCACCGAATAACTATCTCGGAGATACTGCGG 391

QY 597 TGGGCATCTTCTGCTCACTCGATTCCATGTTTCGGAAGTTCGCTCGCATGCTGTGCG 656
DB 392 TCGCAATTTTGTGTTTAACTGGATCCCACTTTTGGTAGTTTTCGCGCAATGCTTTGCG 451

QY 657 TGATCCAAAACAATTCATCCCGGGTACATATTTCAATTTGACGTTTCAATGTTGTGTGA 716
DB 452 TGTTTGAACACAGAAACAGCACCGCGAAAGTAGTCTTCATCGTCACATTCATGTTGTGCGTTA 511
```

## RESULT 6

```
US-10-282-122A-26399
; Sequence 26399, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 26399
; LENGTH: 918
; TYPE: DNA
```

; ORGANISM: Mycobacterium bovis

US-10-282-122A-26399

Query Match 14.8%; Score 191.8; DB 13; Length 918;  
Best Local Similarity 54.4%; Pred. No. 9.3e-53;  
Matches 456; Conservative 0; Mismatches 347; Indels 36;

257	QY	GGGGTGCAGATCTCAAGCTGCCATTGCTGTGGGGATCGGACTGGGGGTCTGTGTTCTT	316
88	Db	 GGCGCCCGTGATTTGCGCGCAGCATCGTGTGGCGCTTTCCATAGSCCTAGTCTCATC	147
317	QY	TTTGGGATTTGCTAAAGCCCATGGGTTGTGTACATCCTGTTGCAGTGTTTATGCTGCA	376
148	Db	 GGGTGCTGTGTTTTCGGCTTCGGGCTTTGGGTTGCCATGTTGGCCGTGCCACCTTGTGC	207
377	QY	GCAACATATGGGAAGTTGGTAGCAGACTTAAAGAGGGCGCTATCATTTGCCACTGCGGATT	436
208	Db	 GCTACCCATAGGTTGTCGGAGGTTGCGGAAGCGGGCTATCTCATCCGGTTATCCG	267
437	QY	ATGATCATCGCGGTGAGGCAATCATCTGCTGTCTATGCGCAATTTGGCACGATGGCAATT	496
268	Db	 TTGCTGATTGCGGGCAGCGCGGTGTGCTGACTGGCGGTTCCGCGCGTCCGCGCA	327
497	QY	TTTGGCGTCTTTTGTGGCCACTGTGTTGGTGTCTGATGATTTTCCGAATTTTACAAATGCG	556
328	Db	 TTGGCGGCTTTGGTGGCATGGTCGTTCTGTGCATGATTTGGGCACTGTTCATGCAGAC	387
557	QY	ACGGAAAAAGAACCCC-----GCAACTATTTTGAGGGACACC	592
388	Db	 AGCGTGCACGCCCGACGACCGCGGTGCACCGTCCGCGGAAACTACTTGTCCGACGTC	447
593	QY	TCTGTGGGCATTTCTGTGCTCACCTGATTTCCATGTTTGGAGCTTTCCTCGGATGCTG	652
448	Db	 TCGGCCACGGTCTTCTGTGGCGGTGTGGTCCCAATGTTCTGCTCTTCGGCGCAATGCTG	507
653	QY	TCGCTGATGCAAAACAATTTCCATCCCGGTACATATTTTCAATTTGACGTTTCATGCTGTGT	712
508	Db	 GTCTACCCGGAAATGGCTCGGATGG-----TGTTCTGCATGATGATCGCG	555
713	QY	GTGATCGCATCGGATGTGGCGGGTATATCGCGGTGTGTTCTTTGGATCGCACCCCAATG	772
556	Db	 GTCACTGCTTCGGATGTGGCGGTGTACGCGGTGGGGGTGCTGTTTGGCAAGCATCCGATG	615
773	QY	GGCGCGTTGGTGTGAGTCGGAAGAAGCTCTTGGGAAGGCTTTTGGCGGCTCCATTCTCTTAGA	832
616	Db	 GTTCCGACGATCAGCCCGAAGAAGTCTGTGGAGGGCTTTTCCGGTTTCGCTGTGTGTCGGG	675
833	QY	TCGGTCACTGGTGCACCTACAGTGTTCATCTCCTGCTCGATCACCACTGGTGTGATGGGTGTG	892
676	Db	 ATCACCGCAACGATCATACCGGACATTTCTTGTGCGCAAAACGCGGTGATTTGGTGCA	735
893	QY	ATCTTGGGTTGTGCCCTAGTTGTGTGGCCACGTTTGGTGACATTGGTTGAGTTCGCAGTTC	952
736	Db	 CTGCTTCGGCGTCTTTTCGTGTCTCACCGCGCTGGGCGACCTGTGTGGAGTTCGCAGGTC	795
953	QY	AAACCGGATTTGGGCATCAAGGATATGTGAAACCTCTTCCAGGCCACGGCGGATTGATG	1012
796	Db	 AAACGTGACCTTCGGCATCAAGACATGGGCGGCTGTACTCCGGCCACCGCGGCTCTGATG	855
1013	QY	GACCGTTTGGATGCTGCTCCCGGCGCGATCGTGACGTGTTGATCTGTAGTGTGAT	1071
856	Db	 GACCGGCTTCGACGCACTATGCTCTTCGGCGGTGGCGGCTTGGATAGTCTCTACACTGCT	914

## RESULT 7

US-10-282-122A-28605  
: Sequence 28605, Application IIS/10282122A

; Sequence 28605, Application US/10282122A

; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

QY 653 TCGCTGATGCAAAACATCCATCCCGGTACATATTTCAATTTGAGCTTCATGCTGTGT 712  
Db 508 GTCTACCCGGAAATGCTCGGATGG-----TGTCTGATGATGATCGCG 555  
QY 713 GTGATCGCATCGGATGCGGGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATG 772  
Db 556 GTCATCGCTTCGATGTCGCGCGGTACCGCGTGGGGGTGCTGTTTGGCAAGCATCCGATG 615  
QY 773 GCGCGGTGGTGAATCCGAAGAGCTTGGGAAGCTTTTCCCGCTCCATTTGCTTAGGA 832  
Db 616 GTTCCGAGCATACCCGGAAGAGTCTGGAGGGCTTTTCCCGTTCGCTGGTGTGCGG 675  
QY 833 TCGGTCACTGTGCACTGAGTTTCACTTCTGCTCGATCACCACACTGGTGTGAGTGTG 892  
Db 676 ATCACCAGCAAGATCATCACCGCACTTTCCTGCTCGCAAAACGCGCTGGATGTTGTCGA 735  
QY 893 ATCTTGGGTTTGGCCCTAGTTTGTGCGCCACGTTGGTGTGACTTGTGATGCGAGTTC 952  
Db 736 CTGCTCGCGGTGCTTTTCTGCTCACCACCGCTGGGACCTGGTGGAGTGGCAGTTC 795  
QY 953 AAACGCGATTTGGGCATCAAGGATATGTCGAACCTTTTCCAGGCCACGCGGATGATG 1012  
Db 796 AAACGTGACCTCGGCATCAAGACATGGGCGCGCTGCTACCCGCCACGCGGCTCTGATG 855  
QY 1013 GACCGTTGGATGATGCTCCCGCGCGGATGTCGAGCTGTTGATCTCTGAGTGTGAT 1071  
Db 856 GACCGGCTCGACGCATACCTGCTTCCCGGTGGCGCTGGATGATGCTCACAACCTGCT 914

RESULT 8  
US-10-282-122A-25992  
; Sequence 25992, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25992  
; LENGTH: 933  
; TYPE: DNA  
; ORGANISM: Mycobacterium avium  
US-10-282-122A-25992  
  
Query Match 13.0%; Score 158.8; DB 13; Length 933;  
Best Local Similarity 52.8%; Pred. No. 4,7e-45;  
Matches 439; Conservative 0; Mismatches 357; Indels 36; Gaps 2;  
  
QY 257 GCGGTCGAGATCTCAAAGCTGCCATTTGCTGTGGGGATCGGACTGGGGTCTCTGTTCTT 316  
Db 103 GCGGAGCTGACCTCGCGGCCGCCATCGCGTGGGCGCGGTATCGGGCCGCTGCTCATC 162  
QY 317 TTGGGAGATTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGACAGTTTATGCTGCA 376  
Db 163 GTACAGCTGTGTTTCCGCGCCCGCTTCTGGGTGCGCATGCTGCGATGGCCCATCTCTGTC 222  
QY 377 GCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436  
Db 223 GCAGCCATGAGTGGTCCGCGGCTGCGGAGGCGGATACGTATCCGGTATTCCG 282  
QY 437 ATGATCATCGGCGGTACAGCAATCATCTGCTGTGTCATGSCCATTTGGCAGATGGCAT 496  
Db 283 TTGCTGCGCGGCGGCGAGCTCAGGTGTGCTGACCTGSCCGTTCCACGCCCGCGCGG 342  
QY 497 TTGGCGTCTTTTGTGGCAGCTGTTGGTGTGATGATATTTCCGAATTTTCTACATATGC 556  
Db 343 TTGGCGGCTTCCGGCTCACCGGTGGCTGCTGTTCTGCGGCTTGTTCATGAGGAC 402  
QY 557 ACGGAAAAAGAGCCCGC-----AACTATTGAGGGACACC 592  
Db 403 AACCGCAAGCGCCCGAGCGCTTCGCCGCTTCGCCCTCGCGCACTACCTGCGCGACGG 462  
QY 593 TCTGTGGCATCTTCTGTCTCATCTGATTCATTTGCGAAGCTTCTGCTGCGATGCTG 652  
Db 463 TCGGCCACCGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522  
QY 653 TCGCTGATGCAAAACAAATTCATCCCGGTACATATTTTCATTTGACGTTTCATGCTGT 712  
Db 523 -----GTCTATCCGCGCGAGGTGCGCGCGGGTGTTCGCTGATGATCACC 570  
QY 713 GTGATCGCATCGGATGCGCGGGTATATCGCGGGTGTGTTCTTTGGATCGCACCAATG 772  
Db 571 GTGGTGGCTCCGACGCTCGCGGCTACGCGGTGCGGTGCTGTTCGCGCAAGCATCCGATG 630  
QY 773 GCGCGCTTGTGAGTCCGAGAGTCTTCGGAAGGCTTTTCGCGGCTCCATTTGCTTTAGGA 832  
Db 631 GTCCCGCGGATCAGCCCCAGAAATCTTGAGAGGCTCTGCGCGCTCTGCTGCTGGGT 890  
QY 833 TCGGTCACTGCTGCTCACTGCTTCTCTCTCGATCACCACATGGTGGATGGGTGTG 892  
Db 691 ATCACCAGCGCCACCTTGGCGCGGACTTTTCTGCGCGCAAGCGCGCTGGGTGGCGG 750  
QY 893 ATCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952  
Db 751 CTGCTGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 810  
QY 953 AAACGCGATTTGGGCATCAAGGATATGTCGAACCTTTCTCCAGGCCACGCGGATGATG 1012  
Db 811 AAGGTGACCTGGGCATCAAGGATGCGCGGCTGCTGCTGCGCGCACGCGCGCTGATG 870  
QY 1013 GACCGTTGATGCGATGCTTCCCGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064  
Db 871 GACCGGCTGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 922

RESULT 9  
US-10-282-122A-27700  
; Sequence 27700, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA 034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 27700  
LENGTH: 939  
TYPE: DNA  
ORGANISM: Mycobacterium leprae  
US-10-282-122A-27700

Query Match 12.6%; Score 164; DB 13; Length 939;  
Best Local Similarity 52.6%; Pred. No. 1.9e-43;  
Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;

QY 257 GCGGTCGAGATCTCAAGCTGCCATTCCTGTGGGATCGGACTGGGGTCTCTGTTCTT 316  
DB 97 GCGGACGTAATCTGCCCGCGCGATCGCGTGGGCTTAAGTATGGTGTCTCGTC 156  
QY 317 TTGGGATGTCTAAGCCCATGGGTGTGATCTCTGTGACAGTCTTATGGCTGCA 376  
DB 157 GCAAGCTGGTGTCTCCGGAATCTGGTGTCTGTGCGTGGCGCAATTTGCTT 216  
QY 377 GCAACATGGGAAGTGTGAGAGACTTAAGAACGGCGCTATCATTTGCCACATGCCGANT 436  
DB 217 GCTAGCCATGAGTGGTGTGCGCGCTACGGAAGCTGGATATGATTCGGGCTATCCCG 276  
QY 437 ATGATCATCGCGGTGAGGCAATCATCTGCTGTCATGTCATGGCCATTTGGACAGTGGCAATT 496  
DB 277 CTGCTCATCGTGGGAGTTTACCGTCTGGTTAACTTGGCCGCTATCGCACCGTGTGFGCA 336  
QY 497 TTGGCGTCTTTTGTGGCACTGTGTGGTGTGATGATTTCCGAATTTT-----CTAC 550  
DB 337 TTGGCTGGCTTCGGTGGCGAGTGGTGTCTGATGATTTGGGGCTTGTGATCGACGAC 396  
QY 551 AATGGCAGGAAAGAACCGCAACTATTGAGGACACCTCTGTGGGCATCTTCGIG 610  
DB 397 AACAGCAGACGACGAATCTCGGAAGCTTTGGCAGTCCCGGCTGCAAAATTACCTG 456  
QY 611 -----CTCACCTGATTCATTTGTTGGAAGCTTCGTCGGATGCTGCTCATGACGAA 665

Db 457 CCGGACCGGTCCGCCACCGTTTTTTTGGCGCGGTGGTCCCGTTGTTGCTCGTTCGCA 516  
QY 666 ACA-----ATTCCATCCCGGTACATATTTTCAATTTTACAGTTTCATGCTGTGTG 715  
Db 517 GCATTTGCTGTTTATCCGAAGATGGTGGCGCGGTATTTCTGCTGATGATTCGGGTG 576  
QY 716 ATCGCATCGGATGTGGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCG 775  
Db 577 GTAGCTCCGATGTGGTGTGTTACACAGTAGGGTACTTTTCGGCAACATCCACTGGTT 636  
QY 776 CCGTTGGTGTCCGAAGAGTCTTGGGAAGCTTTGCCGGCTCCATTTCTTAGGATCG 835  
Db 637 CCAAGGATTAGCCCGAATAAGTCTCGGAGGATTCGCCGTTGCTGTGCTCGCGGACC 696  
QY 836 GTCACCTGGTGCACCTAGTGTTCACCTTCCTGCTCGATCACCACTGGTGGATGGTGTGATC 895  
Db 697 ACCGAACCATCTCTACTCTACCTTCCTGGCTGGCAACACACCGTGGGTGGGCTTTG 756  
QY 896 TTGGGTTGCCCTAGTGTGTGGCCACGTGGGTGACTTGGTTGAGTCGCAAGTTCAA 955  
Db 757 CTCAGTTTTGTACTGGTGTCTCACTGCACGCTGGGGATTTGGTGGAGTCCCAAGTTAA 816  
QY 956 CGCATTTGGGCATCAAGGATATGTCGAACCTTCTTCCAGGCCACGGCGGATTTGATGAC 1015  
Db 817 CGGACCTCGGCATCAAGATATGGCCGGCTGTACCCGGCCACGGTGGCTGATGAC 876  
QY 1016 CGTTTGGATGGCATCTCCGCGCGGATGGTGAAGTGTGATCTGAGTGTGAT 1071  
Db 877 CGCTTGTAGTGTGTGCTGCTGCGGTGTAGCTTGGACCATCTTACGTTGCT 932

RESULT 10  
US-09-864-408A-905  
; Sequence 905, Application US/09864408A  
; Publication No. US2004000947A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shinkets, Richard A.  
; FILE OF INVENTION: No. US2004000947A1  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 905  
; LENGTH: 297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-864-408A-905

Query Match 9.5%; Score 123.4; DB 11; Length 297;  
Best Local Similarity 68.0%; Pred. No. 3.7e-30;  
Matches 172; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 854 GTTCACTTCCTGCTCGATCACCACTGGTGGATGGTGTGATCTTTGGGTGCGCTAGTT 913  
Db 1 GTGCACTTCCTCATCCAGGACCTCTGGTGGATGGGCTGTGCTGGGGGTGCGCTTGC 60  
QY 914 GTGTGGCCACGTGGGTGACTTGGTTGAGTCGCAAGTTCAACCGGATTTGGGATCAAG 973  
Db 61 CTCTCGCCACCATGGCGGACCTCTGTGAGAGGACAGTTCAAGCGCGAGCTGGGAATCAAG 120  
QY 974 GATATGTGGAACCTTCTCCAGGCCACCGCGGATTTGATGGACCGTTTGGATGGCATGCTC 1033  
Db 121 GACATGTGAACTTCTCTCCGGCCACCGCGGGCTGATGGACCCCTTGGACGGGATGCTG 180  
QY 1034 CCGGCGCGGATGGTGAAGTGTGATCTGATGTGATCAGACGCTGTATTCCTCGTAA 1093  
Db 181 CCGCGCGGACGCGGACCTCATGCTCAACGCGCGGCGGATCATGAACCCGCTGTAG 240

QY 1094 AGCTTGGCCAGC 1106  
Db 241 CGGCCCTGGCAGC 253

RESULT 11

US-09-738-626-2215/c  
; Sequence 2215, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent In ver. 3.0  
; SEQ ID NO 2215  
; LENGTH: 258  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-2215

Query Match 9.0%; Score 117; DB 9; Length 258;  
Best Local Similarity 100.0%; Pred. No. 4.7e-28;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1184 AAGGCTTTACGGATTTTCTTAAGTGGCGAGCGTACTCAACATGGCGCAGCCACCAAC 1243  
Db 258 AAGGCTTTACGGATTTTCTTAAGTGGCGAGCGTACTCAACATGGCGCAGCCACCAAC 199  
QY 1244 AAGGCCATATCAATGACACCGGTAATGGCTGCTAGTAGGAAACCGATTCCGGCTGG 1300  
Db 198 AAGGCCATATCAATGACACCGGTAATGGCTGCTAGTAGGAAACCGATTCCGGCTGG 142

RESULT 12

US-10-156-761-2610  
; Sequence 2610, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 2610

; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1176)  
US-10-156-761-2610

Query Match 8.5%; Score 110.2; DB 15; Length 1176;  
Best Local Similarity 54.6%; Pred. No. 2.3e-25;  
Matches 242; Conservative 0; Mismatches 198; Indels 3; Gaps 1;  
QY 624 CATTTGTCGGAAGCTTCGCTGCGATGCTGTGCTGATGCAAAACAATTCATCCCGGTA 683  
Db 713 CGTTCTAGTCCCGTTCTCTGGCAGCTTCGTGCGATGATGCTACGGCCGACGAGCTC 772  
QY 684 CATATTTCAATTTGACGTTCACTGCTGTGTGATCGCATCGGATGCGGCGGTATATCG 743  
Db 773 CGCGCGGGTGTCTCAGTTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 832  
QY 744 CGGGTGTGTCTTTTGGATCGCACCAATGGCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 803  
Db 833 TCGGCTGGCGTTTCGGCAAGCACAGCTCGCCCGCGCATCAGCCCGGCAAGACCCGCG 892  
QY 804 AAGGCTTTCGGCTCCATTGTCTTAGATCGTCACTGGTCACTCAGTGTCTCACTTCC 863  
Db 893 AGGGCTGTGTGCGAGCGGTCTCGTTGCGATGTTGGCGGCGCGCTGTGCAATGGAGTTCC 952  
QY 864 TGCTCGATCACCA---CTGGTGGATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 920  
Db 953 TGATCGACGACGGCAGCTGTGTGCGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1012  
QY 921 CCAGTGTGGTGACTGT 980  
Db 1013 CCAGCTCGCGACCTCGCGAGTCCATGATCAAGCGGATCTGGGCATCAAGGACATGG 1072  
QY 981 CGAACCTTCTTCAGGCCACGCGGATTCATGACCGTTTGTGATGCGATGCTCCCGGCG 1040  
Db 1073 GCAGCTGTGTGCGGGCACGGCGCATGATGACCGGTGTGTGTGTGTGTGTGTGTGTGT 1132  
QY 1041 CGATGTTGACGT 1063  
Db 1133 CTCCGGT 1155

RESULT 13

US-10-156-761-1/c  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4187715)



OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Query Match 8.5%; Score 110.2; DB 15; Length 9025608;  
Best Local Similarity 54.6%; Pred. No. 5.5e-23;  
Matches 242; Conservative 0; Mismatches 198; Indels 3; Gaps 1;  
  
QY 624 CATTTGTCGGAAGCTTCGGTCGGATGCTGCTCGTCATGCAAAACAAATTCATCCCGGTA 683  
DB 3223095 CGTTCTACGTCCTCGGTCCTGGGACAGTTCGTCGCGATGATCTCAGGCCACAGACGGTC 3223036  
  
QY 684 CATATTTCATTGACGTTTCATGCTGTGTGTGATCGCATCGATGATGGCGGGTATATCG 743  
DB 3223035 CGCGCGGGTGTCTACGTTCTCTGCTGCTCACCGTGTGAGCGACACGGCGCGTACGGCA 3222976  
  
QY 744 CGGGTGTGTTCTTCGATGCGACCAATAGGCGCGCTGCTGAGTCCGAGAGTCTTGG 803  
DB 3222975 TCGGCTGGGTTTCGCGAAGCACAGGCTCGCCCGCGCATCAGCCCGCGAAGACCGCG 3222916  
  
QY 804 AAGGCTTTCCGGCTCCATGTGTTTAGGATCGGTCACTGTGTCACACTCAGTGTTCACITTC 863  
DB 3222915 AGGGCTGTGTCGAGCGGCTCTCGTTGCGATGTTGGCGCGCGCTGTGATGAGTTCC 3222856  
  
QY 864 TGCTCGATCACCA---CTGTGATGAGTGTGATCTTGGTGTGCTTGGCTAGTGTGTGG 920  
DB 3222855 TGATCGACACGGCAGCTGTGTGGCAGGCGCTGCTCGGCTTCGCGGTGCGCCCGCAGTG 3222796  
  
QY 921 CCACGTTGGGTGACATTGTTGAGTCGCGATTCAACACGGATTTCGGCATCAAGGATATGT 980  
DB 3222795 CCACGCTCGGACCTCGCGAGTCCATGATCAAGCGGATCTGGGATCAAGGACATGG 3222736  
  
QY 981 CGAACCTTCTTCAGGCACGCGGATGATGAGACGTTTGGATGGCATGCTCCCGCGCG 1040  
DB 3222735 GCACGCTGTGCGCGGCGACGCGCATGAGACGCGCTGGAATCGTGTGCGCAGCG 3222676  
  
QY 1041 CGATGCTGACGTTGTTGATCTG 1063  
DB 3222675 CTCGGTGTGTTGTTGCTG 3222653

RESULT 14  
US-10-470-565-1  
; Sequence 1, Application US/10470565  
; Publication No. US20040126870A1  
; GENERAL INFORMATION:  
; APPLICANT: Societe des Produits Nestle S.A.  
; TITLE OF INVENTION: NCC2705 - the genome of a Bifidobacterium  
; FILE REFERENCE: 80290/WO  
; CURRENT APPLICATION NUMBER: US/10/470,565  
; PRIORITY FILING DATE: 2003-07-29  
; PRIOR APPLICATION NUMBER: EP 01102050.0  
; PRIOR FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 2256646  
; TYPE: DNA  
; ORGANISM: Bifidobacterium longum  
US-10-470-565-1  
  
Query Match 5.4%; Score 69.8; DB 17; Length 2256646;  
Best Local Similarity 52.0%; Pred. No. 7.9e-10;  
Matches 192; Conservative 0; Mismatches 162; Indels 15; Gaps 1;  
  
QY 721 ATCCGATGTCGGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGCGCGGT 780  
DB 1884103 ATCCGATGTCGGCGGTGTTGTTCCGTCGCTCGGCAACACAAAGCTCAGCCCGG 1884162  
  
QY 781 GGTGAGTCCGAAGAAGCTTTGGGAAGGCTTTGGCGGCTCATTGTCTTAGGATCGGTAC 840  
DB 1884163 CATCTCGCGAAGAAGCTTTGGAGGCTTTGCGGCTCGATCTCTTCGCCATGCTG 1884222  
  
QY 841 TGGTGCACTCAGTGTTCCTACTCTCTGCTGATCAC-----CACTGGTGGAT 885

DB 1884223 CGCATTTGCCGTTATTCGATGCGATGATGATGCTCGAATGGCCACCTGTTGGTGGT 1884282  
QY 886 GGGTGTGATCTTCGGTGTGCTGCTAGTGTGTGCGCCACGTTGGGTGACTTGGTTGAGTC 945  
DB 1884283 GCCGATCGTGGCGGTATCTCATCGTCCGCGTGGGACGTTTCGGTGAATTGTGTGCTTC 1884342  
QY 946 GCAGTTCAACCGGATTTGGGATCAAGGATATGTGCAACCTTCTTCCAGGCCACGGCGG 1005  
DB 1884343 CATGCTCAACCGGATTCGGTATCAAGACATGCGCCATCTGCTCAAGGGGACGGCGG 1884402  
QY 1006 ATTGATGACCGGTTTGGATGCGATGCTCCCGCGCGATGCTGCTGATGCTGATCTGAT 1065  
DB 1884403 TGTGATGATGCTGTGATGATTCATTTTGTGATGATGATGATGATGATGATGATG 1884462  
QY 1066 TGTGATCAG 1074  
DB 1884463 GATCACCGG 1884471

RESULT 15  
US-10-282-122A-33608  
; Sequence 33608, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 33608  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Pseudomonas syringae  
US-10-282-122A-33608

Query Match 5.3%; Score 68.4; DB 13; Length 813;  
Best Local Similarity 51.2%; Pred. No. 1.8e-11;  
Matches 212; Conservative 0; Mismatches 196; Indels 6; Gaps 2;



```
Qy 640 CGCTGCGATGCTGCTGATGCAABACAATTCATCCCGGTACATATTTTCATTTTGAC 699
Db |||
Qy 360 CGCGTGGCAGGGGTGCTGCTCATCAAGCAATGGCCGCTGGGCAACTGGCTGATCCTGTC 419
Db |||
Qy 700 GTTCATGCTGTGTGTGATCGCATCGGATGTGGGCGGTATATCCCGGGTGTTCCTTTGG 759
Db |||
Qy 420 GGTGATGGTGTGTGTGGCGGTGACATCGGCGGTATTTCTCCGCAAGGCTTTCGG 479
Db |||
Qy 760 ATCCACCCCAATGGCGCGTGTGGTCACTCCGAGAAAGTCTTGGGAAGGCTTTCGCGCTC 819
Db |||
Qy 480 CAGCGCAAGCTGGCTCCCAAGSTCACTCCCGGTAAAGCTGGGAGGCGGTATACGCGG 539
Db |||
Qy 820 CATTGTCTTAGGATCGG---TCACTGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCA 876
Db |||
Qy 540 TCTGCTGGTCAGTCTGGGTATCACCGCGGCAGTCGGTGTGGCGTGACTGGACTGTCT 599
Db |||
Qy 877 CTGCTGATGGGTGTGATCTTGGGTGTGGCTTGTGCTAGTGTGTGCG---CCAGCTTGGTGA 933
Db |||
Qy 600 TCAGTTCATCGCCGCACTGTTGGGTGCGCTGTGATGCTGTTCATTTCCGTGATCGGTGA 659
Db |||
Qy 934 CTTGTTGAGTCGCAGTTCAAACCGGATTTGGGCGATCAAGGATATGTGGAACCTTCTTCC 993
Db |||
Qy 660 CCTCACCGAAGCATGTTCAAGCCCGAGTCCGGGTCAAGGACAGCAGTATCTGCTGCC 719
Db |||
Qy 994 AGGCCACGGCGGATGATGGACCGTTTGGATGGCATGCTCCCGCGCGGATGGT 1047
Db |||
Qy 720 TGGGCACGGCGGCTTCTGGACCGCATCGACAGCTGACGCGCGCCCATTCGGT 773
Db |||
```

Search completed: August 17, 2004, 21:10:38  
Job time : 886 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:45:38 ; Search time 785 Seconds

(without alignments)  
7035.232 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagtccttctgcgaagga.....aggaaacgattccgctgg 1300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*

- 1: geneseq1980s.\*
- 2: geneseq1990s.\*
- 3: geneseq2000s.\*
- 4: geneseq2001as.\*
- 5: geneseq2001bs.\*
- 6: geneseq2002s.\*
- 7: geneseq2003as.\*
- 8: geneseq2003bs.\*
- 9: geneseq2003cs.\*
- 10: geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	1300	6	Aa168796 C. glutam
C 2	1300	100.0	34980	5	Aah68531 C glutami
C 3	1300	100.0	34980	5	Aah68530 C glutami
4	978	75.2	978	4	Aaf67792 Coryneb
5	978	75.2	978	4	Aaf67792 Coryneb
6	855	65.8	855	5	Aaf67181 C glutami
7	342.2	26.3	876	7	Aca29916 Coryneb
8	191.8	14.8	918	7	Aca38529 Prokaryot
9	191.8	14.8	921	7	Aca40735 Prokaryot
C 10	191.8	14.8	110000	4	Continuation (32 o
C 11	191.8	14.8	110000	4	Continuation (32 o
12	168.8	13.0	933	7	Aca38122 Prokaryot
13	164	12.6	939	7	Aca39830 Prokaryot
14	123.4	9.5	297	6	Ahn75506 Human syn
C 15	117	9.0	258	5	Aah67180 C glutami
C 16	117	9.0	369	7	Aca00855 C. gluram
17	110.6	8.5	1019	4	Aas59898 Propionib
18	110.6	8.5	1019	4	Aas59898 Propionib
19	69.8	5.4	34980	6	Abq81849 Bifidobac
20	68.4	5.3	813	7	ACA45738 Prokaryot
21	67.2	5.2	813	7	ACA43503 Prokaryot
22	64	4.9	849	7	ACf69725 Photorhab
C 23	64	4.9	110000	7	Continuation (28 o

24	64	4.9	110000	7	ACF65386_1
C 25	62.4	4.8	19547	4	AAS59601_
C 26	62.4	4.8	19547	7	ACf64530 Propionib
27	59	4.5	843	7	ACA53364 Prokaryot
28	58.8	4.5	936	4	AAS54139 Pseudomon
29	58.8	4.5	936	7	ACA42333 Prokaryot
30	58.4	4.5	2017	6	AAL40183 Isoprenoi
31	58.2	4.5	789	7	ACA35433 Prokaryot
32	58.2	4.5	795	7	ACA37071 Prokaryot
33	56.6	4.4	819	7	ACA23571 Prokaryot
34	55.8	4.3	855	7	ACA27424 Prokaryot
35	55.2	4.2	789	7	ACA36374 Prokaryot
36	55.2	4.2	1733	6	ABQ70612 Listeria
37	55.2	4.2	110000	6	Continuation (14 o
38	55	4.2	819	7	ACA26133 Prokaryot
39	54.6	4.2	748	7	ACA48951 Prokaryot
40	54.6	4.2	858	7	ACA51226 Prokaryot
41	53.2	4.1	846	7	ACA38990 Prokaryot
42	53.2	4.1	945	7	ACF70555 Photorhab
43	53.2	4.1	99629	4	Aaf28550 Genomic f
44	53.2	4.1	110000	7	Continuation (37 o
C 45	53.2	4.1	110000	7	Continuation (12 o

ALIGNMENTS

RESULT 1  
Aa168796  
ID Aa168796 standard; DNA; 1300 BP.  
XX  
AC Aa168796;  
XX  
DT 22-JAN-2002 (first entry)  
XX  
DE C. glutamicum ATCC 13032 cdsA DNA.  
XX  
KW Coryneform bacterium; cdsA; phosphatide-cytidyltransferase;  
KW L-lysine production; amino acid production; medicine; animal nutrition;  
KW growth rate; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
FH Key Location/Qualifiers  
RBS 187..195  
FT /\*tag= a  
FT CDS 200..1093  
FT /\*tag= b  
FT /\*product= "cdsA"  
XX  
XX DE10021828-A1.  
XX  
PD 08-NOV-2001.  
XX  
PF 04-MAY-2000; 2000DE-01021828.  
XX  
PR 04-MAY-2000; 2000DE-01021828.  
XX  
PA (DEGS ) DEGUSSA AG.  
XX (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.  
XX  
XX Nampoothiri M, Moeckel B, Pfeifferle W, Eggeling L, Sahn H;  
XX WPI; 2002-018672/03.  
XX P-PSDB; AAG80226.  
XX  
XX New mutant coryneform bacterium, useful for production of amino acids,  
XX especially lysine, has increased activity of phosphatide-cytidyl-  
XX transferase.  
XX  
XX Claim 9; Page 10-12; 16pp; German.  
XX  
XX This invention describes a novel genetically modified coryneform

CC bacterium (A) in which activity of the cdsA gene, encoding phosphatide-  
 CC cytidyl transferase, is increased. (A) are used for fermentative  
 CC production of amino acids, especially L-lysine, useful in human medicine,  
 CC animal nutrition and pharmaceuticals. Nucleic acid, or fragments, derived  
 CC from the cdsA gene are used (i) as primers for polymerase chain reaction  
 CC of cdsA genes or (ii) as probes for isolating cDNA and genes that are  
 CC highly homologous with cdsA. Overexpression of the cdsA gene results in  
 CC increased production of amino acids, also more rapid growth to higher  
 CC cell densities. This sequence encodes the Corynebacterium glutamicum ATCC  
 CC 13032 phosphatide-cytidyl transferase (cdsA) described in the method  
 CC of the invention  
 XX  
 SQ

Sequence 1300 BP; 286 A; 289 C; 371 G; 354 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1300; DB 6; Length 1300;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTTGCTCGCAAGAAAGAAAGACTGATGAGGTCTAGAGACCTTTATGCAATG 60  
 DB 1 GAAGTCGTTGCTCGCAAGAAAGAAAGACTGATGAGGTCTAGAGACCTTTATGCAATG 60  
 QY 61 GCTTAAGAAATTACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTGCTT 120  
 DB 61 GCTTAAGAAATTACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTGCTT 120  
 QY 121 GCAAGAGTGTGGAGAAATTTTTCGAAATGCTGGCACCATCAACAGTGAATGTTA 180  
 DB 121 GCAAGAGTGTGGAGAAATTTTTCGAAATGCTGGCACCATCAACAGTGAATGTTA 180  
 QY 181 GAAACTTCAAGAGAACCCATGATGAAACCGGAGCAATCACCCTGTCATGAGATGCC 240  
 DB 181 GAAACTTCAAGAGAACCCATGATGAAACCGGAGCAATCACCCTGTCATGAGATGCC 240  
 QY 241 CAAACCCAAAAATAATCGGGTTCGAGATCTCAAAGCTGCCATTTGCTGGGATCGGACT 300  
 DB 241 CAAACCCAAAAATAATCGGGTTCGAGATCTCAAAGCTGCCATTTGCTGGGATCGGACT 300  
 QY 301 GGGGTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360  
 DB 301 GGGGTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360  
 QY 361 AGGTTTATGCTGCGCAACATGGAAGTGGTAGCAGACTTAAAGAAAGCGGCTATCA 420  
 DB 361 AGGTTTATGCTGCGCAACATGGAAGTGGTAGCAGACTTAAAGAAAGCGGCTATCA 420  
 QY 421 TTGCGACTGCGGATATGATCATCGGGTTCAGCAATCATCTGGCTGATGCGCAT 480  
 DB 421 TTGCGACTGCGGATATGATCATCGGGTTCAGCAATCATCTGGCTGATGCGCAT 480  
 QY 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGCCACTGTTGGTGTGCTGATGATTTCCG 540  
 DB 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGCCACTGTTGGTGTGCTGATGATTTCCG 540  
 QY 541 AATTTTCTAATGCGCAAGAAAGAAAGCGCCCAACTATTTTGAAGGACACCTCTGTTGG 600  
 DB 541 AATTTTCTAATGCGCAAGAAAGAAAGCGCCCAACTATTTTGAAGGACACCTCTGTTGG 600  
 QY 601 CATCTTGTGCTCAGCTGATTTCCATTTCTCGGAAGCTTCCCTGCGATGCTGCTGAT 660  
 DB 601 CATCTTGTGCTCAGCTGATTTCCATTTCTCGGAAGCTTCCCTGCGATGCTGCTGAT 660  
 QY 661 GCAAAACAAATTCATCCCGGTTACATTTTCAATTTGAGCTTCATGCTGTGTGATCGC 720  
 DB 661 GCAAAACAAATTCATCCCGGTTACATTTTCAATTTGAGCTTCATGCTGTGTGATCGC 720  
 QY 721 ATCGATGCGGGGATATCGGGGTTCTTTTGGATCGCACCAATGGCGGCTT 780  
 DB 721 ATCGATGCGGGGATATCGGGGTTCTTTTGGATCGCACCAATGGCGGCTT 780  
 QY 781 GGTGAGTCCGAAGAGTCTTGGGAAGGCTTTTGGCGGCTCCATTTGCTAGATCGGTAC 840  
 DB 781 GGTGAGTCCGAAGAGTCTTGGGAAGGCTTTTGGCGGCTCCATTTGCTAGATCGGTAC 840

QY 841 TGTGCACTCAGTGTTCACCTTCTCTGCTCGATCACCACCTGCTGATGGTGTGATCTTGG 900  
 DB 841 TGTGCACTCAGTGTTCACCTTCTCTGCTCGATCACCACCTGCTGATGGTGTGATCTTGG 900  
 QY 901 TTGTGCCCTAGTGTGTGGCCACAGTTGGGTGACCTTGGTGTGATGCTGACCTTCAAACGCGA 960  
 DB 901 TTGTGCCCTAGTGTGTGGCCACAGTTGGGTGACCTTGGTGTGATGCTGACCTTCAAACGCGA 960  
 QY 961 TTGTGGGATCAAGAGATATGTCGAACCTTTCTTCAGGCCACGCGGATTCATGACCGTTT 1020  
 DB 961 TTGTGGGATCAAGAGATATGTCGAACCTTTCTTCAGGCCACGCGGATTCATGACCGTTT 1020  
 QY 1021 GGATGCGATGCTCCCGCGCGGATGATGACGTTGATCTCTGATGCTGATCAGCAGCTC 1080  
 DB 1021 GGATGCGATGCTCCCGCGCGGATGATGACGTTGATCTCTGATGCTGATCAGCAGCTC 1080  
 QY 1081 GTATCCGTGTAAAGCTTGGGCCAGCTTTTAAAGTTCAAAAACCTTGAAGGCCCTGAGGTG 1140  
 DB 1081 GTATCCGTGTAAAGCTTGGGCCAGCTTTTAAAGTTCAAAAACCTTGAAGGCCCTGAGGTG 1140  
 QY 1141 CATAACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGATTT 1200  
 DB 1141 CATAACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGATTT 1200  
 QY 1201 TTCTTAACCTGGGACCGTACTCAAACATCGGCACGCGCCACCAAGCGCCATTAATCATG 1260  
 DB 1201 TTCTTAACCTGGGACCGTACTCAAACATCGGCACGCGCCACCAAGCGCCATTAATCATG 1260  
 QY 1261 CACCGGTAAATGGCTGTAGTAAACCGATTTCCGGCTGG 1300  
 DB 1261 CACCGGTAAATGGCTGTAGTAAACCGATTTCCGGCTGG 1300

RESULT 2

AAH68531/c  
 ID AAH68531 standard; DNA; 349980 BP.

XX AC AAH68531;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 7066.

XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX KW organic acid synthesis; ds.

XX OS Corynebacterium glutamicum.

XX PN BP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 99JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX PA (KYOW ) KYOWA HAKKO KOGYO KK.

XX FI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX FI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX DR WPI; 2001-376931/40.

XX PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene.  
 XX Disclosure; SEQ ID NO 7066; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Corynebacterium, and identifying a homologue of a gene derived from  
CC corynebacterium. Corynebacterium bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 349980 BP: 80724 A; 98367 C; 90490 G; 80399 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1300; DB 5; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAAGTCGTTGCTCGAAGGAAAGAACTGATGAGGCTTAGAGACCTTTATCGCAATG 60  
DB 37324 GAAGTCGTTGCTCGAAGGAAAGAACTGATGAGGCTTAGAGACCTTTATCGCAATG 37265  
QY 61 GCCTAAGAAATACGCTCTTCAACGTCGATTGGCGGGGAAAGACGCTTTCTTTTGCTT 120  
DB 37264 GCCTAAGAAATACGCTCTTCAACGTCGATTGGCGGGGAAAGACGCTTTCTTTTGCTT 37205  
QY 121 GCAAGAGTGTTGGAAGAAATTTTTCGAAAATGTTGGCACCATCAACAGTGACATTGTTA 180  
DB 37204 GCAAGAGTGTTGGAAGAAATTTTTCGAAAATGTTGGCACCATCAACAGTGACATTGTTA 37145  
QY 181 GAAACTTCAAGGAGAACCCATGATGACCGGAGCAACATCACCCTCCATGAGGATGCC 240  
DB 37144 GAAACTTCAAGGAGAACCCATGATGACCGGAGCAACATCACCCTCCATGAGGATGCC 37085  
QY 241 CAACCCCAAAATTAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTGGGATCGGAAT 300  
DB 37084 CAACCCCAAAATTAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTGGGATCGGAAT 37025  
QY 301 GGGGTCCTGTTCTTTGGGGATGTCCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 360  
DB 37024 GGGGTCCTGTTCTTTGGGGATGTCCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 36965  
QY 361 AGGTTTATGCTGCAGCAACATGGAAATGGTAGCAGACTTAAAGAGCGGCTATCA 420  
DB 36964 AGGTTTATGCTGCAGCAACATGGAAATGGTAGCAGACTTAAAGAGCGGCTATCA 36905  
QY 421 TTGCGACTGCGGATATGATCATCGCGGTGAGCAATCATCTGGCTGTCATGCCCAT 480  
DB 36904 TTGCGACTGCGGATATGATCATCGCGGTGAGCAATCATCTGGCTGTCATGCCCAT 36845  
QY 481 TGGCAGGATGGCATTTTGGGCTCTTTTGTGCGCACTGTGTTGGTGTGATGATTTCCG 540  
DB 36844 TGGCAGGATGGCATTTTGGGCTCTTTTGTGCGCACTGTGTTGGTGTGATGATTTCCG 36785  
QY 541 AATTTTCTACATGCGACGGAAGAAAGACCGCACTATTATGAGGACACCTCTGTGGG 600  
DB 36784 AATTTTCTACATGCGACGGAAGAAAGACCGCACTATTATGAGGACACCTCTGTGGG 36725  
QY 601 CATCTCTGTCCTCACTGATTCCTTGTGGAAGCTTCGCTCGATGCTGCTGAT 660  
DB 36724 CATCTCTGTCCTCACTGATTCCTTGTGGAAGCTTCGCTCGATGCTGCTGAT 36665  
QY 661 GCAAAACAAATTCCTCCGGGTACATATTTTATTTGACGTTTCATGCTGTGTGTGATCGC 720  
DB 36664 GCAAAACAAATTCCTCCGGGTACATATTTTATTTGACGTTTCATGCTGTGTGTGATCGC 36605  
QY 721 ATCGGATGGCGGGATATATCGCGGTGTGTTCTTTGATGCGACCAATGCGGCCCTT 780  
DB 36604 ATCGGATGGCGGGATATATCGCGGTGTGTTCTTTGATGCGACCAATGCGGCCCTT 36545  
QY 781 GGTGAGTCGGAAGAGCTCTTGGAGGCTTTCGCGGCTCCATGCTTAGGATCGGTAC 840  
DB 36544 GGTGAGTCGGAAGAGCTCTTGGAGGCTTTCGCGGCTCCATGCTTAGGATCGGTAC 36485

QY 841 TGGTGCACCTCAGTGTTCATCTTCCTGCTCGATCACCACCTGGTGGATGGTGTGATCTTGGG 900  
DB 36484 TGGTGCACCTCAGTGTTCATCTTCCTGCTCGATCACCACCTGGTGGATGGTGTGATCTTGGG 36425  
QY 901 TTGTGCCCTAGTGTGTGTCGCCACAGTTTGGGTGACCTTGGTTCAGTCGCGAGTTTCAACGCGA 960  
DB 36424 TTGTGCCCTAGTGTGTGTCGCCACAGTTTGGGTGACCTTGGTTCAGTCGCGAGTTTCAACGCGA 36365  
QY 961 TTTGGGCAATCAAGGATATGTGCGAACCTTCTTCAGGCCACCGCGGATTTGATGACCGCTTT 1020  
DB 36364 TTTGGGCAATCAAGGATATGTGCGAACCTTCTTCAGGCCACCGCGGATTTGATGACCGCTTT 36305  
QY 1021 GGATGCGATGCTCCCGCGCGGATGCTGACGPGTGTGATCCTGAGTGTGATCAGCAGCTC 1080  
DB 36304 GGATGCGATGCTCCCGCGCGGATGCTGACGPGTGTGATCCTGAGTGTGATCAGCAGCTC 36245  
QY 1081 GTATCGCTGTAAGCTTTGGCCAGCTTTAAAGTTCAAAAGCTTGAAGGGCTCTGAGTG 1140  
DB 36244 GTATCGCTGTAAGCTTTGGCCAGCTTTAAAGTTCAAAAGCTTGAAGGGCTCTGAGTG 36185  
QY 1141 CATACGCTGTCACCTCAGCGGCTTTTGGCTGTCAAAAGTTTAAAGGGCTTTTACGATTT 1200  
DB 36184 CATACGCTGTCACCTCAGCGGCTTTTGGCTGTCAAAAGTTTAAAGGGCTTTTACGATTT 36125  
QY 1201 TTCTTAACCTGGCGAGCGTACTCAAAACATGCGCACGCCACCAACAGCGCCATAATCAATG 1260  
DB 36124 TTCTTAACCTGGCGAGCGTACTCAAAACATGCGCACGCCACCAACAGCGCCATAATCAATG 36065  
QY 1261 CACCGGTAATGCTGCTAGTAGGAAACCGATTCGCGCTGG 1300  
DB 36064 CACCGGTAATGCTGCTAGTAGGAAACCGATTCGCGCTGG 36025  
RESULT 3  
AAH68530/c  
ID AAH68530 standard; DNA; 349980 BP.  
XX  
AC AAH68530;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum coding sequence fragment SEQ ID NO: 7065.  
XX  
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;  
XX  
KW organic acid synthesis; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN EP1108790-A2.  
XX  
PD 20-JUN-2001.  
XX  
PF 18-DEC-2000; 2000EP-00127688.  
XX  
PR 16-DEC-1999; 99JP-00377484.  
XX  
PR 07-APR-2000; 2000JP-00159162.  
XX  
PR 03-AUG-2000; 2000JP-00280988.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
XX  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX  
DR WPI; 2001-376931/40.  
XX  
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
XX  
PT mutation point of a gene, measuring expression of a gene, analyzing  
XX  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX  
PS Disclosure; SEQ ID NO 7065; 246pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of nucleotide and protein  
CC

CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and analysing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from  
 CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the European Patent Office  
 XX  
 SQ Sequence 349980 BP; 86896 A; 98023 C; 80939 G; 84122 T; 0 U; 0 Other;

Query Match 100.0%; Score 1300; DB 5; Length 349980;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

337324 GAAAGTCGTTGCTCGCAAGGAAAGAACTGATGGAGGCTTAGAGACCTTTATCGCAATG 337265  
 DB  
 61 GGTAAAGATTAACGTCGTTTCAACCTGATTTGGCGGGGAAACGACGCTTTCTTTTGCTT 120  
 DB GCTTAAAGATTAACGTCGTTTCAACCTGATTTGGCGGGGAAACGACGCTTTCTTTTGCTT 337205  
 121 GCAAGAGTGTGTTGGAAGATTTTTCGAAATGCTGGCACCATTCAACAGTACATTTGTTA 180  
 DB GCAAGAGTGTGTTGGAAGATTTTTCGAAATGCTGGCACCATTCAACAGTACATTTGTTA 337145  
 181 GAAACTTCAAGAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCC 240  
 DB GAAACTTCAAGAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCC 337085  
 241 CAAACCCAAAATAATGCGGTGAGATCTCAAAGCTGCCATTTGCTGTGGGATCGGACT 300  
 DB CAAACCCAAAATAATGCGGTGAGATCTCAAAGCTGCCATTTGCTGTGGGATCGGACT 337025  
 301 GGGGTCCTGCTGTTCTTTTGGGGATTTGCTTAAGCCCATGGGGTTGTACATCTCTGTTGC 360  
 DB GGGGTCCTGCTGTTCTTTTGGGGATTTGCTTAAGCCCATGGGGTTGTACATCTCTGTTGC 336965  
 361 AGTTTATAGCTGCGAGCAACATGGAAGTTGTTAGCAGACTTAAAGAGCGCGCTATCA 420  
 DB AGTTTATAGCTGCGAGCAACATGGAAGTTGTTAGCAGACTTAAAGAGCGCGCTATCA 336905  
 421 TTGCGCACTGCGGATTAATGATCATCGCGGTGAGCAATCATCTGCTGTCATGGCCATT 480  
 DB TTGCGCACTGCGGATTAATGATCATCGCGGTGAGCAATCATCTGCTGTCATGGCCATT 336845  
 481 TGGCAGCATGGGCATTTTGGCGTCTTTTGGCCACTGTGTGGTGTGATGATTTCCG 540  
 DB TGGCAGCATGGGCATTTTGGCGTCTTTTGGCCACTGTGTGGTGTGATGATTTCCG 336785  
 541 AATTTCTACAAATGCGCAGGAAAGAGCCGCACTATTGAGGACACCTCTGTGGG 600  
 DB AATTTCTACAAATGCGCAGGAAAGAGCCGCACTATTGAGGACACCTCTGTGGG 336725  
 601 CATCTTGTGCTCACTGGATTTCCATTTGTTGCGGAGCTTCGCTGCGATGCTGTGCTGAT 660  
 DB CATCTTGTGCTCACTGGATTTCCATTTGTTGCGGAGCTTCGCTGCGATGCTGTGCTGAT 336685  
 661 GCAAAACAATTCATCCCGGTGATATTTTCAATTTGACGTTCAATGCTGTGTGATTCG 720  
 DB GCAAAACAATTCATCCCGGTGATATTTTCAATTTGACGTTCAATGCTGTGTGATTCG 336605  
 721 ATCCGATGTGGCGGTATATTCGGGTGTGTTCTTTTGGATGCAACCAATGCGCGGTT 780  
 DB ATCCGATGTGGCGGTATATTCGGGTGTGTTCTTTTGGATGCAACCAATGCGCGGTT 336545  
 781 GGTGAGTCCGAAGAGTCTTTGGGAGGCTTTTGGCGCTCCATTGCTTTAGGATTCGCTCAC 840  
 DB GGTGAGTCCGAAGAGTCTTTGGGAGGCTTTTGGCGCTCCATTGCTTTAGGATTCGCTCAC 336485

QY 841 TGGTGCACTCAGTGTTCACCTTCTCTGCTCGATCACCACCTGGTGGATGATCTTGGG 900  
 DB TGGTGCACTCAGTGTTCACCTTCTCTGCTCGATCACCACCTGGTGGATGATCTTGGG 336484  
 QY 901 TTTGGCCCTAGTGTGTGGGCCACGTTGGGTGAGTTCGAGTTCGAGTTCAAAACGCGA 960  
 DB TTTGGCCCTAGTGTGTGGGCCACGTTGGGTGAGTTCGAGTTCGAGTTCAAAACGCGA 336424  
 QY 961 TTTGGGCATCAAGGATATGTCGAACCTTCTTCAGGCCACGCGGATTCGATGACCGTTT 1020  
 DB TTTGGGCATCAAGGATATGTCGAACCTTCTTCAGGCCACGCGGATTCGATGACCGTTT 336364  
 QY 1021 GGATGCACTCTCCCGCCCGCATGCTGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 1080  
 DB GGATGCACTCTCCCGCCCGCATGCTGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 336304  
 QY 1081 GTATCCGTGCTAAGCTTGGGCCAGCTTTAAGTTCAAAAACCTTGAAGCGCTGAGGTG 1140  
 DB GTATCCGTGCTAAGCTTGGGCCAGCTTTAAGTTCAAAAACCTTGAAGCGCTGAGGTG 336244  
 QY 1141 CATACGCTGTGCACCTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGATTT 1200  
 DB CATACGCTGTGCACCTCAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGATTT 336184  
 QY 1201 TTCTTAACCTGGGACGCTGACTCAAAACATCGGACGCCCAACCAAGCGCCATAATG 1260  
 DB TTCTTAACCTGGGACGCTGACTCAAAACATCGGACGCCCAACCAAGCGCCATAATG 336124  
 QY 1261 CACCGGTAATGCTGCTAGTAGGAACCGATTCGCGCTGG 1300  
 DB CACCGGTAATGCTGCTAGTAGGAACCGATTCGCGCTGG 336064  
 RESULT 4  
 ID AAF67792 standard; DNA; 978 BP.  
 XX AAF67792;  
 AC AAF67792;  
 XX  
 DT 11-APR-2001 (first entry)  
 XX  
 DE  
 XX  
 KW Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:99.  
 KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;  
 KW membrane construction and membrane transport protein; petroleum spill;  
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;  
 KW identification; microorganism; fine chemical production; transfection;  
 KW genome mapping; genetic engineering; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 XX WO200100805-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 XX PF 23-JUN-2000; 2000WO-IB000926.  
 XX  
 PR 25-JUN-1999; 99US-0141031P.  
 PR 08-JUL-1999; 99DE-01031454.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031563.  
 PR 09-JUL-1999; 99DE-01032122.  
 PR 09-JUL-1999; 99DE-01032124.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032128.  
 PR 09-JUL-1999; 99DE-01032180.  
 PR 09-JUL-1999; 99DE-01032182.  
 PR 09-JUL-1999; 99DE-01032190.  
 PR 09-JUL-1999; 99DE-01032191.  
 PR 09-JUL-1999; 99DE-01032209.  
 PR 09-JUL-1999; 99DE-01032212.  
 PR 09-JUL-1999; 99DE-01032227.

PR	09-JUL-1999	99DE-1032228.
PR	09-JUL-1999	99DE-1032228.
PR	09-JUL-1999	99DE-1032230.
PR	14-JUL-1999	99DE-1032927.
PR	14-JUL-1999	99DE-1033005.
PR	14-JUL-1999	99DE-1033006.
PR	27-AUG-1999	99DE-1040764.
PR	27-AUG-1999	99DE-1040765.
PR	27-AUG-1999	99DE-1040766.
PR	27-AUG-1999	99DE-1040830.
PR	27-AUG-1999	99DE-1040831.
PR	27-AUG-1999	99DE-1040832.
PR	27-AUG-1999	99DE-1040833.
PR	31-AUG-1999	99DE-1041378.
PR	31-AUG-1999	99DE-1041379.
PR	31-AUG-1999	99DE-1041395.
PR	03-SEP-1999	99DE-1042077.
PR	03-SEP-1999	99DE-1042078.
PR	03-SEP-1999	99DE-1042079.
PR	03-SEP-1999	99DE-1042088.
XX		
PA	(BADI ) BASF AG.	

XX  
PI Dommeius M. Kroeger B. Schroeder H. Zelder O. Haberhauer G;

XX  
DR WPI; 2001-071486/08.  
DR P-PSDB: AAB76559.

XX Corynebacterium glutamicum nucleic acids encoding membrane construction  
PT and membrane transport proteins or their portions, useful for typing or  
PT identifying C. glutamicum or related bacteria, and as markers for  
PT transformation.

XX  
PS  
Claim 3: page 322-323; 1119pp; English; English

AAAF67743 to AAFA68080 encode the *Corynebacterium glutamicum* membrane construction and membrane transport (MCT) proteins given in AAB76510 to AAB76847. The MCT nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in *C. glutamicum* or related bacteria (e.g. *Brevibacterium lactofermentum*), the typing or identification of *C. glutamicum* or related bacteria, as reference points for mapping *C. glutamicum* genome, and as markers for transformation. AAF68082 and AAF68082 represent sequencing primers which are used in an example from the present invention

XX  
SO  
Sequence 978 BP: 200 A: 220 C: 285 G: 273 T: 0 U: 0 Other:

Query Match 75.2%; Score 978; DB 4; Length 978;  
Best Local Similarity 100.0%; Pred. No. 3.e-300;  
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	136	AGAA	TTTTTTT	CGAAAAT	GTCTGC	ACCAT	CAAC	CAGT	GACAT	TGTT	AGAAA	CTT	CAGG	AGA	195						
DB	1	AGAA	TTTTTTT	CGAAAAT	GTCTGC	ACCAT	CAAC	CAGT	GACAT	TGTT	AGAAA	CTT	CAGG	AGA	60						
QY	136	ACCC	ATGAT	GAAC	CGGAG	CAAC	AT	CAAC	CGGT	CCAT	GAGAT	GCC	CAAA	ACCC	AAAAA	255					
DB	61	ACCC	ATGAT	GAAC	CGGAG	CAAC	AT	CAAC	CGGT	CCAT	GAGAT	GCC	CAAC	ACCC	AAAAA	120					
QY	256	TCCG	GGTC	GAGAT	CTT	CAAA	GCT	GCAT	TCT	GTG	GGG	GAT	CGA	CT	GGGGT	CT	GGT	CT	315		
DB	121	TCCG	GGTC	GAGAT	CTT	CAAA	GCT	GCAT	TCT	GTG	GGG	GAT	CGA	CT	GGGGT	CT	GGT	CT	180		
QY	316	TTT	GGG	GA	TG	TCT	CA	AAG	CCCAT	TGGG	TGG	TAT	GC	AT	CCCT	CG	TTT	TAT	TGG	CT	375
DB	181	TTT	GGG	GA	TG	TCT	CA	AAG	CCCAT	TGGG	TGG	TAT	GC	AT	CCCT	CG	TTT	TAT	TGG	CT	240
QY	376	AGCA	CAT	CGG	AA	GT	TGG	TAG	CGA	CT	TAA	AGA	AGG	GGC	TAT	CA	TTT	TGC	CA	CT	435
DB	241	AGCA	CAT	CGG	AA	GT	TGG	TAG	CGA	CT	TAA	AGA	AGG	GGC	TAT	CA	TTT	TGC	CA	CT	300
QY	436	TAT	GAT	CAT	CGG	CGG	T	CAGG	CA	TAT	CT	TGG	CT	TGT	CAT	GCC	CA	TTT	TGG	CA	495

01-JUL-1999; 99DE-01030476.  
 02-JUL-1999; 99US-0142101P.  
 03-JUL-1999; 99DE-01031415.  
 04-JUL-1999; 99DE-01031418.  
 05-JUL-1999; 99DE-01031419.  
 06-JUL-1999; 99DE-01031420.  
 07-JUL-1999; 99DE-01031424.  
 08-JUL-1999; 99DE-01031428.  
 09-JUL-1999; 99DE-01031434.  
 10-JUL-1999; 99DE-01031435.  
 11-JUL-1999; 99DE-01031443.  
 12-JUL-1999; 99DE-01031453.  
 13-JUL-1999; 99DE-01031457.  
 14-JUL-1999; 99DE-01031465.  
 15-JUL-1999; 99DE-01031478.  
 16-JUL-1999; 99DE-01031510.  
 17-JUL-1999; 99DE-01031541.  
 18-JUL-1999; 99DE-01031573.  
 19-JUL-1999; 99DE-01031592.  
 20-JUL-1999; 99DE-01031632.  
 21-JUL-1999; 99DE-01031634.  
 22-JUL-1999; 99DE-01031636.  
 23-JUL-1999; 99DE-01032125.  
 24-JUL-1999; 99DE-01032126.  
 25-JUL-1999; 99DE-01032130.  
 26-JUL-1999; 99DE-01032186.  
 27-JUL-1999; 99DE-01032206.  
 28-JUL-1999; 99DE-01032227.  
 29-JUL-1999; 99DE-01032228.  
 30-JUL-1999; 99DE-01032229.  
 31-JUL-1999; 99DE-01032230.  
 01-AUG-1999; 99DE-01032922.  
 02-AUG-1999; 99DE-01032926.  
 03-AUG-1999; 99DE-01032928.  
 04-AUG-1999; 99DE-01033004.  
 05-AUG-1999; 99DE-01033005.  
 06-AUG-1999; 99DE-01033006.  
 07-AUG-1999; 99US-0148613P.  
 08-AUG-1999; 99DE-01040764.  
 09-AUG-1999; 99DE-01040765.  
 10-AUG-1999; 99DE-01040766.  
 11-AUG-1999; 99DE-01040832.  
 12-AUG-1999; 99DE-01041378.  
 13-AUG-1999; 99DE-01041379.  
 14-AUG-1999; 99DE-01041380.  
 15-AUG-1999; 99DE-01041394.  
 16-AUG-1999; 99DE-01041396.  
 17-AUG-1999; 99DE-01042076.  
 18-AUG-1999; 99DE-01042077.  
 19-AUG-1999; 99DE-01042079.  
 20-AUG-1999; 99DE-01042086.  
 21-AUG-1999; 99DE-01042087.  
 22-AUG-1999; 99DE-01042088.  
 23-AUG-1999; 99DE-01042095.  
 24-AUG-1999; 99DE-01042124.  
 25-AUG-1999; 99DE-01042129.  
 26-AUG-1999; 99DE-01042129.  
 27-AUG-1999; 2000US-0187970P.  
 (BADI ) BASF AG.  
 Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberthauer G;  
 WPI; 2001-137957/14.  
 P-PSDB; AAB80192.  
 Nucleic acids from *Corynebacterium glutamicum* encoding metabolic pathway  
 proteins, useful for producing fine chemicals in microorganisms,  
 including organic acids, nonproteinogenic amino acids, and purine and  
 pyrimidine bases.  
 Claim 3; Page 1674; 1737pp; English.  
 AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic

CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP  
 CC nucleic acids are useful for the production of fine chemicals in  
 CC microorganisms, including organic acids, nonproteinogenic amino acids,  
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated  
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,  
 CC vitamins, cofactors, polyketides and enzymes  
 XX  
 SQ Sequence 978 BP; 200 A; 220 C; 285 G; 273 T; 0 U; 0 Other;  
 Query Match 75.2%; Score 978; DB 4; Length 978;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-300;  
 Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 136 AGAATTTTTCGAATGCTGCACCATCAACAGTCAGATTTTGTAGAACTTCACAGGAGA 195  
 Db 1 AGAATTTTTCGAATGCTGCACCATCAACAGTCAGATTTTGTAGAACTTCACAGGAGA 60  
 QY 196 ACCATGAATGAACCGAGCAACATCACCGTCCATGAGGATGCCCAACCAAAATAA 255  
 Db 61 ACCATGAATGAACCGAGCAACATCACCGTCCATGAGGATGCCCAACCAAAATAA 120  
 QY 256 TCGGGTCGAGATCTCAAAGCTGCGATTGCTGTGGGATCGGACTGGGGTCTCTGGTCT 315  
 Db 121 TCGGGTCGAGATCTCAAAGCTGCGATTGCTGTGGGATCGGACTGGGGTCTCTGGTCT 180  
 QY 316 TTTGGGATTGTCTTAAGCCCATGGGTTGGTACATCTCTGTCAGGTTTATGCTGTC 375  
 Db 181 TTTGGGATTGTCTTAAGCCCATGGGTTGGTACATCTCTGTCAGGTTTATGCTGTC 240  
 QY 376 AGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCATCCGAT 435  
 Db 241 AGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCATCCGAT 300  
 QY 436 TATGATCATCGCGCTCAGGCAATCATCTGCTGTGTCATGTCGATGATTTCCGAATTTCTACAATGG 555  
 Db 301 TATGATCATCGCGCTCAGGCAATCATCTGCTGTGTCATGTCGATGATTTCCGAATTTCTACAATGG 420  
 QY 496 TTTGGGCTTTTGTGGCCACTGTGTTGTCGATGATTTCCGAATTTCTACAATGG 555  
 Db 361 TTTGGGCTTTTGTGGCCACTGTGTTGTCGATGATTTCCGAATTTCTACAATGG 420  
 QY 556 CACGAAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGGATCTTCGTGCTCAC 615  
 Db 421 CACGAAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGGATCTTCGTGCTCAC 480  
 QY 616 CTGGATTCCATTGTTGGGAAGCTTCGCTGCCATGTCGCTGATGCAAAAACAAATTCAT 675  
 Db 481 CTGGATTCCATTGTTGGGAAGCTTCGCTGCCATGTCGCTGATGCAAAAACAAATTCAT 540  
 QY 676 CCGGGTACATATTTTCATTTTGAAGCTTCATGCTGTGTGATGTCGATGCGGATGGGCGG 735  
 Db 541 CCGGGTACATATTTTCATTTTGAAGCTTCATGCTGTGTGATGTCGATGCGGATGGGCGG 600  
 QY 736 GTATATCGCGGTGTGTTCTTTTGGATCGCACCCCAATGGCGCGTGTGTGATGCGGAGAA 795  
 Db 601 GTATATCGCGGTGTGTTCTTTTGGATCGCACCCCAATGGCGCGTGTGTGATGCGGAGAA 660  
 QY 796 GTCTTGGGAAGGCTTTCGCGCTCCATTGCTTAGGATCGGTCATCTGTGCTCAGTGT 855  
 Db 661 GTCTTGGGAAGGCTTTCGCGCTCCATTGCTTAGGATCGGTCATCTGTGCTCAGTGT 720  
 QY 856 TCATTCCTGCTCGATCACCACCTGGTGGATGGGTGTGATCTTGGGTTGTGCGCTAGTGT 915  
 Db 721 TCATTCCTGCTCGATCACCACCTGGTGGATGGGTGTGATCTTGGGTTGTGCGCTAGTGT 780  
 QY 916 GTGGCCACGTTGGGTGACATTTGTTGAGTGCAGTTCAAACCGCATTTGGGATCAAGGA 975  
 Db 781 GTGGCCACGTTGGGTGACATTTGTTGAGTGCAGTTCAAACCGCATTTGGGATCAAGGA 840  
 QY 976 TATGTGGAACCTTCTTCAGGCCACCGGATGATGACCGTGTGATGCGATGCTCCC 1035  
 Db 841 TATGTGGAACCTTCTTCAGGCCACCGGATGATGACCGTGTGATGCGATGCTCCC 900



QY 1036 GGCGCGATGTCAGCTGGTGTGATCTCTGAGTGTGATCAGCAGCTCGTATCCCGTCGTAAG 1095  
Db 901 GGCGCGATGTCAGCTGGTGTGATCTCTGAGTGTGATCAGCAGCTCGTATCCCGTCGTAAG 960  
QY 1096 CTTGGGCCAGCTTTAAAGT 1113  
Db 961 CTTGGGCCAGCTTTAAAGT 978  
RESULT 6  
AAH67181  
ID AAH67181 standard; DNA; 855 BP.  
XX AC AAH67181;  
XX 26-SEP-2001 (first entry)  
XX C glutamicum coding sequence fragment SEQ ID NO: 2216.  
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
KW organic acid synthesis; ds.  
XX Corynebacterium glutamicum.  
OS  
XX EPI108790-A2.  
XX 20-JUN-2001.  
XX 18-DEC-2000; 2000EP-00127688.  
XX 16-DEC-1999; 99JP-00377484.  
PR 07-APR-2000; 2000JP-00159162.  
PR 03-AUG-2000; 2000JP-00280988.  
XX (KYOW) KYOWA HAKKO KOGYO KK.  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai X, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.  
XX P-PSDB; AAG91962.  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.  
XX Claim 8; SEQ ID NO 2216; 246pp + Sequence Listing; English.  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Coryneform bacterium, and identifying a homologue of a gene derived from  
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a nucleic acid described  
CC in the exemplification of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from the European Patent Office  
XX  
SQ Sequence 855 BP; 159 A; 194 C; 257 G; 245 T; 0 U; 0 Other;  
Query Match 65.8%; Score 855; DB 5; Length 855;  
Best Local Similarity 100.0%; Pred. No. 4.2e-261;  
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 236 ATGCCCAACCCAAAATAATCGGGTCCGAGATCTCAAGCTGCCATTCGCTGGGGATC 295  
Db 1 ATGCCCAACCCAAAATAATCGGGTCCGAGATCTCAAGCTGCCATTCGCTGGGGATC 60  
QY 296 GGACTGGGGGCTCGTGTCTTTTGGGGATGTGCTTAAGCCCATGGGTGGTACATCCTC 355

Db 61 GGACTGGGGGCTCGTGTCTTTTGGGGATGTGCTTAAGCCCATGGGTGGTACATCCTC 120  
QY 356 GTTGACAGTCTTTATGCTGTGACGCAACATGGCAAGCTTGGTAGCAGACTTAAAGAGGGGCG 415  
Db 121 GTTGACAGTCTTTATGCTGTGACGCAACATGGCAAGCTTGGTAGCAGACTTAAAGAGGGGCG 180  
QY 416 TATCATTTGCCACTGCCGATTAATGATCATCGGGGTGAGGCAATCATCTGCTGTCAATGG 475  
Db 181 TATCATTTGCCACTGCCGATTAATGATCATCGGGGTGAGGCAATCATCTGCTGTCAATGG 240  
QY 476 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGGCCACTGTGTGCTGCTGATGAT 535  
Db 241 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGGCCACTGTGTGCTGCTGATGAT 300  
QY 536 TTCCGAATTTTCAATGGCAACGAAAAAGAGCCCGCAACTATTGAGGAGACACCTCT 595  
Db 301 TTCCGAATTTTCAATGGCAACGAAAAAGAGCCCGCAACTATTGAGGAGACACCTCT 360  
QY 596 GTGGGCATCTTCGTGCTCAGCTGATTCATTTGCGAAGCTTCGTCGCGATGCTGTGCG 655  
Db 361 GTGGGCATCTTCGTGCTCAGCTGATTCATTTGCGAAGCTTCGTCGCGATGCTGTGCG 420  
QY 656 CTGATGCAAAAACAATTCATCCCGGTACATATTTTCACTTTTACCTTCATGCTGTGTG 715  
Db 421 CTGATGCAAAAACAATTCATCCCGGTACATATTTTCACTTTTACCTTCATGCTGTGTG 480  
QY 716 ATCGCATCGGATGTGGGGGGTATATCGGGGTGTGTTCTTTGGATCGCACCCATGGCG 775  
Db 481 ATCGCATCGGATGTGGGGGGTATATCGGGGTGTGTTCTTTGGATCGCACCCATGGCG 540  
QY 776 CCGTTGGTGAGTCCGAAAGTCTTTGGGAAGGCTTTGCGGCTCCATTTGTTAGGATCG 835  
Db 541 CCGTTGGTGAGTCCGAAAGTCTTTGGGAAGGCTTTGCGGCTCCATTTGTTAGGATCG 600  
QY 836 GTCACTGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCACTGGTGGATGGTGTGATC 895  
Db 601 GTCACTGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCACTGGTGGATGGTGTGATC 660  
QY 896 TTGGGTGTGCGCCTAGTGTGTCGCGCCAGTGGGTGAGCTTGGTGGTGGTGGTGGTGGT 955  
Db 661 TTGGGTGTGCGCCTAGTGTGTCGCGCCAGTGGGTGAGCTTGGTGGTGGTGGTGGTGGT 720  
QY 956 CGCGATTTGGGCGATCAAGGATATGTGAAACCTTCTTCCAGGCCACGGCGGATGATGGAC 1015  
Db 721 CGCGATTTGGGCGATCAAGGATATGTGAAACCTTCTTCCAGGCCACGGCGGATGATGGAC 780  
QY 1016 CGTTTGGATGGCATGCTCCCGGCGCGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1075  
Db 781 CGTTTGGATGGCATGCTCCCGGCGCGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840  
QY 1076 AGCTCGTATCCGTCG 1090  
Db 841 AGCTCGTATCCGTCG 855  
RESULT 7  
ACA29916  
ID ACA29916 standard; DNA; 876 BP.  
XX AC ACA29916;  
XX 19-JUN-2003 (first entry)  
XX Prokaryotic essential gene #11573.  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX Corynebacterium diphtheriae.  
XX WO200271183-A2.  
XX 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.  
 XX PF 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;  
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX DR P-PSDB; ABU26046.  
 XX DR WPI: 2003-029926/02.  
 XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX PS Claim 14; SEQ ID NO 17786; 1766pp; English.  
 XX CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: the sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 876 BP; 171 A; 184 C; 244 G; 277 T; 0 U; 0 Other;  
 Query Match 25.3%; Score 342.2; DB 7; Length 876;  
 Best Local Similarity 63.1%; Pred. No. 8.1e-98;  
 Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;  
 QY 237 TGCCCAACCCAAATAATCGGGTGCAGATCTCAAGCTGCCATTCGTGCGGATCG 296  
 DB 32 TGCTTAACCCAGAGAAATCGGCTGTGCGAACTGAAGCCGCGATCAGCGTTGGAAATG 91  
 QY 297 GACTGGGGTCTCGTCTTTTGGGAGATTCCTTAAGCCCATCGGGTGTACATCCTCG 356  
 DB 92 GCTTGGCGGCCCTCGTATTCGCTATCTTTGTTATTCCTTTTGGTGTGATCCATTAG 151  
 QY 357 TTGCAGGTTTTATGGCTGACAGCACTGGAAGTTGGTAGCAGACTTAAAGAGCGGGCT 416  
 DB 152 TGCGTATCGCAATCGCACTGGAAGTGGAGAGTAGAGCGCCGTTTATAGAGCCGGTT 211  
 QY 417 ATCATTTGCCACTGCCGATTAATGATCATCGGGGTGAGGCAATCATCTGCTGTCTATGCG 476  
 DB 212 ATCTTTTGCACAGTTGGGTTCATGCTTTATTTGGGGGACAAGTCATGTTTGTGGCTTAGTTGGC 271  
 QY 477 CATTTGGCAGATGGGCATTTTGGCGCTCTTTTGTGCCACTGTGTTGGTGTGCTGATGTAAT 536  
 DB 272 CATTTGGTCCGAGGAGTAGTGGCTGGTTCGTCGGCGTGTACTCTCGTAGATGTTTG 331  
 QY 537 TCCGAATTTTCTACATGGCAGCGAAAGAGCCGCAACTATTTAGAGGACACCTCTG 596  
 DB 332 GTCCGGCTTTTCCATCATGCTGCTTCTATGCCACCGGAAACTATCTGCAGATCTGCCG 391  
 QY 597 TGGGCATCTTCGTCGCTCACTGGATTCGATTCGAGAGCTTCGCTCGAGTCTGCTGCG 656  
 DB 392 TCGCAATTTTGTGTTTAACTGGATCCCACTTTTGGTAGTTTTCGGGCAATGCTTTGCG 451  
 QY 657 TGATGCAAAACAATTCATCCCGGGTACATATTTTTCAGCTTTCATGCTGTGTGTA 716  
 DB 452 TGTTTGAACAGAAACAGCACCGGAAAGTACTTTCATGTCACATTCATGTTTGTGGTTA 511  
 QY 717 TCGCATCGGATCGGGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCGC 776  
 DB 512 TCGCCTCTCGAGTAGCGGTATATTTGACAGGCGTGTATGCGATCGCACCCCATGGCAC 571  
 QY 777 CTTTGGTGTAGTCGGAAGAGTCTTTGGGAAGGCTTTTCCGGGCTCCATTGCTTTAGGATCGG 836  
 DB 572 CTGCGTAAAGCCCAAGAAATCGTGGAGGGCTTTTATCGGTTCTGTAGTCTTTGSCATGA 631  
 QY 837 TCACTGGTCACTAGTGTTCATCTCTGCTCGATCACACACTGGTGGATGGGTGATCT 896  
 DB 632 TTGTTGGAGCATTCATCTGTGCTCTATTTATTAGGCGCATCAATGGTGGTGGGATTAGTTC 691  
 QY 897 TGGGTGTGCCCTAGTGTGTCGCGCACCGTTGGGTGACTTGGTTGAGTCGAGTTCAGAAC 956  
 DB 692 TCGGATTTGGACTTGTGATCTGCGCGACACTGGGCGATCTTGTGCAATCCAGTTTAA 751  
 QY 957 GCGATTTGGGCATCAAGGATATGTGCAACCTTTTCCAGGCCACGGCGGATTTGATGGACC 1016  
 DB 752 GGGAACTCGGATCAAGACATGTCGCAATTTTGGCCGCGCATGGTGTCTGATGGATC 811  
 QY 1017 GTTTCGATGGCATGCTCCCGCGCGAGTGTGAGTGGTTCATCTGAGTGTGAT 1071  
 DB 812 GCGTTGACGCTATGTTGCCATCGGCTATGAGTGTGATCTTGTAGTGTGAT 866  
 RESULT 8  
 ACA38529  
 ID ACA38529 standard; DNA; 918 BP.  
 XX ACA38529;  
 AC ACA38529;  
 DT 19-JUN-2003 (first entry)  
 XX Prokaryotic essential gene #20186.  
 DE Prokaryotic essential gene #20186.  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX Mycobacterium bovis.  
 XX WO200277183-A2.  
 PD 03-OCT-2002.  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU34659.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 26399; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcr\_sequences  
XX  
XX Sequence 918 BP; 128 A; 278 C; 314 G; 198 T; 0 U; 0 Other;  
Query Match 14.8%; Score 191.8; DB 7; Length 918;  
Best Local Similarity 54.4%; Pred. No. 6.5e-50;  
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;  
257 GCGGTCGAGATCTCAAGTCGCCATTCCTGTGGGATCGGACTGGGGTCTGTGTTCTT 316  
88 GCGGCGCGTGTATTTGCGGCGAGCGATCGTGTGGGCTTTCCATAGGCGCTAGTCTCATC 147  
317 TTGGGATGTCCTAAGCCATCGGTTGGTGTACATCTCGTTGCGAGTTTATGGCTGCA 376  
148 GCGGTGCTGGTGTTCGTTCCGCGGTTTGGTTGCCATCGTGGCGTCGCCACCTTGGTC 207  
377 GCACATCGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436  
208 GCTACCCATGAGTGTGCGGAGGTTCGGGAAGCGGCTATCTCATCCCGTTATCCCG 267  
437 ATGATCATCGGGTCAGGCAATCATCTGGCTGTATCGGCTATGGCAGCATGGGATT 496  
268 TTGCTGATTGGCGGCGAGCCCGCGTGTGGCTGACCTGGCGGCTTCGGCGGTCGGCGCA 327  
497 TTGGCGTCTTTTGTGGCCACTGTGTGTGTGTGTATTTCCGAATTTTCTACAATGGC 556  
328 TTGGCGGGCTTTGGTGGATGTGCTGTCTGTGATGATTGGGACTGTTTCATGAGGAC 387  
557 ACGGAAAAAGAACGCC-----GCACTATTTCAGGGACACC 592

388 AGCGTAGCGCCGCGACGACGCGCGGTGCACCGTTCGCGGGAACACTACTTGTCCGACGTC 447  
593 TCTGTGGGCATCTTCGTCCTCACCTGGATTCCATGTTCGGAAGCTTCGTCGCAATGCTG 652  
448 TCGGCCACGGTCTTCTTGGCGGTGTGGGTCCCATTTGTTCTCTCTTTCGGGCGCAATGCTG 507  
653 TCGCTGATGCAAAACAATTCCATCCCGGTACATATTTTCACTTTTGGAGCTTCATGCTGTGT 712  
508 GTCATCCCGGAAATGGCTCGGGATGGG-----TGTTCTGCATGATGATCGCG 555  
713 GTGATCGCATCGGATGTGGCGGGTATATCGCGGTGTGTTCCTTTGGATCGCACCCCAATG 772  
556 GTCATCGCTTCCGATGTGCGCGCTACGCCGTGGGGTGTGTTTGGCAAGCATCCGATG 615  
773 GCGCGCTTGTGAGTCCGAAAGTCTTGGAAAGCTTTGCGGCTCCATTTGCTTAGGA 832  
616 GTTCCGACGATCAGCCCAAGAGTGTGGGAGGCTTTGCGGGTTCCTGCTGTGCGGG 675  
833 TCGGTCACTGTGTGCACTCAGTGTTCACCTCTCTGCTCGATCACCACCTGGTGGATGGTGTG 892  
676 ATCACCAGCAATCATCACCGGACTTCTTGGTGGCAAAACGCGCTGGATTGGTGA 735  
893 APTCTGGGTGTGCGCTAGTGTGTGCGCCACGTTGGGTGACTGTGTTGAGTCCGAGTTC 952  
736 CTGCTCGCGGTCTTTTCTGTGTGTACCAACCGCTGGCGGACCTGTGTGAGTTCGAGTTC 795  
953 AAACCGGATTTGGGATCAAGATATGCGAACCTTCTCCAGGCGACGCGGATTGATG 1012  
796 AAACGTGACCTCGGATCAAGACATGGCGCGCTGTACCCGCGCACGCGGCTGTGATG 855  
1013 GACCGTGTGGATGGATGCTCCCGCGCGATGTGCGTGTGATGCTGTGATGCTGTGATG 1071  
856 GACCGGCTCGAGCGCATACTGCTTCCGCGGTGGCGCTGGATAGTCTCCTCACACTGCT 914  
RESULT 9  
ACA40735  
ID ACA40735 standard; DNA; 921 BP.  
AC ACA40735;  
XX 19-JUN-2003 (first entry)  
DE Prokaryotic essential gene #22392.  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
XX Mycobacterium tuberculosis.  
XX WO200277183-A2.  
XX 03-OCT-2002.  
XX 21-MAR-2002; 2002WO-US009107.  
XX 21-MAR-2001; 2001US-00815242.  
XX 06-SEP-2001; 2001US-00948993.  
XX 25-OCT-2001; 2001US-0342923P.  
XX 08-FEB-2002; 2002US-00072851.  
XX 08-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU36865.  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.



AAI99683_04	WP	400001	510000
AAI99683_05	WP	500001	610000
AAI99683_06	WP	600001	710000
AAI99683_07	WP	700001	810000
AAI99683_08	WP	800001	910000
AAI99683_09	WP	900001	1010000
AAI99683_10	WP	1000001	1110000
AAI99683_11	WP	1100001	1210000
AAI99683_12	WP	1200001	1310000
AAI99683_13	WP	1300001	1410000
AAI99683_14	WP	1400001	1510000
AAI99683_15	WP	1500001	1610000
AAI99683_16	WP	1600001	1710000
AAI99683_17	WP	1700001	1810000
AAI99683_18	WP	1800001	1910000
AAI99683_19	WP	1900001	2010000
AAI99683_20	WP	2000001	2110000
AAI99683_21	WP	2100001	2210000
AAI99683_22	WP	2200001	2310000
AAI99683_23	WP	2300001	2410000
AAI99683_24	WP	2400001	2510000
AAI99683_25	WP	2500001	2610000
AAI99683_26	WP	2600001	2710000
AAI99683_27	WP	2700001	2810000
AAI99683_28	WP	2800001	2910000
AAI99683_29	WP	2900001	3010000
AAI99683_30	WP	3000001	3110000
AAI99683_31	WP	3100001	3210000
AAI99683_32	WP	3200001	3310000
AAI99683_33	WP	3300001	3410000
AAI99683_34	WP	3400001	3510000
AAI99683_35	WP	3500001	3610000
AAI99683_36	WP	3600001	3710000
AAI99683_37	WP	3700001	3810000
AAI99683_38	WP	3800001	3910000
AAI99683_39	WP	3900001	4010000
AAI99683_40	WP	4000001	4110000
AAI99683_41	WP	4100001	4210000
AAI99683_42	WP	4200001	4310000
AAI99683_43	WP	4300001	4403765

Query Match 14.8%; Score 191.8; DB 4; Length 110000;  
Best Local Similarity 54.4%; Pred. No. 9.4e-49;  
Matches 456; Conservative 0; Mismatches 347; Indels 36;

257	QY	GGGGTCGAGATCTCAAAAGCTGCCATTCCTCTGGGATCGCACTGGGGGTCCTGGTCTT	316
85640	Db	CGCCGCCGTGATTTTCGGCGACGATCGTGTGGCCCTTCCATAGGCCCTAGTCCTCATC	85581
317	QY	TTTGGGATGTGCTTAAGCCCATGGGTGGTACATCTCGTTCGAGGTTTATGGCTGCA	376
85580	Db	GGGGTCTGTGTTTCGTTCCGGCGTTTGGTTGCCATCTGTGGCCGTCGCCACCTTGTC	85521
377	QY	GCAACATGGGAAGTTGGTAGCAGACTTAAAGAAGCGGCTATCATTTGCCACTGCGGATT	436
85520	Db	GCTACCCATGAGTGTGCGGAGTTGCGGAAAGCGGCTATCTCATCCCGTTATCCGC	85461
437	QY	ATGATCATCGGCGGTACGCAATCATCTGGCTGTCTATGCCCATTTGGCAGCATGGCAATT	496
85460	Db	TTGCTGATTTGGCGGCGAGCCCGGTGTGGTACCTCGCCCGTTCGGCGCGGTCCGGCA	85401
497	QY	TTGGCGTCTTTTGTGGCACTGTGTTGGTGCATGTATTCCGAATTTTCTACAAATGGC	556
85400	Db	TTGGCGGGCTTTGGTGGCATGTCGTTGTCTGCATGATTTGGCGACTGTTTATGCAGAC	85341
557	QY	ACGGAAAAAGAACCC-----GCAACTATTTCAGGGACACC	592
85340	Db	AGCGTACGCGGCCGACGACCGCGGTGCACCGTCGCGGGAAACTACTTTCGACGTC	85281
593	QY	TCGTGGGCATCTTCGTGCTACCTGGATTCAATTGTTGGAGAGTTTCGCTGCGATGCTG	652
85280	Db	TCGGCCACGGCTCTTCCTGCGCCGTGGGTCCCAATGTCTGCTCTTCGGGCAATGCTG	85221

RESULT 11		AAI199683_11/C		Continuation (32 of 44) of AAI199683 from base 3100001 (Mycobacterium tuberculosis strain H37Rv)	
WP	Sequence	Split into 44 fragments	LOCUS AAI199683	Accession	AAI19683
WP	Fragment Name	Begin	End		
WP	AAI199683_00	1	110000		
WP	AAI199683_01	100001	210000		
WP	AAI199683_02	200001	310000		
WP	AAI199683_03	300001	410000		

QY 653 TCCTGATGCAAAACAAATTCATCCCGGTACATATTTTCATTTGACGTTTCATGCTGTGT 712  
 Db 85220 GTCTACCCGGAATAAGCTCGGATGGG-----TGTTCTGATGATGATCGCG 85173  
 QY 713 GTATCGCATCGGATGCGCGGATATATCGCGGTGTGTTCTTTGGATCGCACCAATG 772  
 Db 85172 GTATCGCTTCGATGTCGGCGGTACGCGGTGGGGTCTGTTTGGCAAGCATCCGATG 85113  
 QY 773 GCGCCGTTGTGATGTCGAAGAAGCTTTGGGAAGCTTTGCGGTCATGTTCTTTAGGA 832  
 Db 85112 GTTCCACCATCAGCCCGAAGAAGCTGTCGGAGGGCTTTGCGGTTCTGCTGTGCGGG 85053  
 QY 833 TCGGTCACTGGTCACTGATGTTCACTTCTCTGCTCGATCACCATCGTGGATGGTGTG 892  
 Db 85052 ATCACCGCAACATCATCACCGGACTTCTCTGTCGCGACCTGGCGACCTGGTAGTGGTGA 84993  
 QY 893 ATCTGGGTTGTCCTAGTTGTGTCGCGCACCTGTTGGTGAATGTTGAGTCGCGATTC 952  
 Db 84992 CTGCTCGCGGTGCTTTTCGTCTCACACCGCGCTGGCGACCTGGTAGTTCGCGATG 84933  
 QY 953 AAACGGATTTGGCATCAGGATATGTCGAACCTTCTTCCAGGCCACGGCGGATGATG 1012  
 Db 84932 AAACGTGACTCGGCATCAACACATGGGCGCCCTGCTACCGGCCACGCGGCTCTGATG 84873  
 QY 1013 GACCGTTTGGATGATGCTCCCGCGCGATGTCAGTGTGATGCTGATGCTGATG 1071  
 Db 84872 GACCGGCTCGAGCATATGCTCTTCCGCGGTGGCGGCTGGATGCTCACACTGCT 84814

## RESULT 12

ACA38122  
 ID ACA38122 standard; DNA; 933 BP.

XX ACA38122;

XX ACA38122;

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #19779.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

KW Mycobacterium avium.

XX WO20027183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

PI WPI: 2003-029926/02.

XX P-PSDB; ABU34252.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 25992; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC the gene product or that has an activity against a biological pathway;  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 933 BP; 102 A; 318 C; 343 G; 170 T; 0 U; 0 Other;

Query Match 13.0%; Score 168.8; DB 7; Length 933;

Best Local Similarity 52.8%; Pred. No. 1.4e-42;

Matches 439; Conservative 0; Mismatches 357; Indels 36; Gaps 2;

QY 257 GCGGTTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATCGGACTGGGGTCTGTTCTT 316  
 Db 103 GCGGACGTGACCTCGGGGCCCATCGCGGTGGCGCGGTATCGGCGGTGCTATC 162  
 QY 317 TTGGGATGTCCTAAGCCCATGGGGTTGGTACATCTCTGTTCAGGTTTATGGCTGCA 376  
 Db 163 GTACAGCTGGTGTTCGCGCCCGCTTCTGGGTGCGATCGTCGATGCCCATCTCGTGC 222  
 QY 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436  
 Db 223 GCAGCCATGAGTGTGTCGCGCGCTGCGGAGCGGATAGCTATCCGGTATTCCG 282  
 QY 437 ATGATCATCGCGGTGAGCAATCATCTGGGTGTATGCGCATTTGGCAGTGGCAATT 496  
 Db 283 TTGTGCGCGCGCGCAGCTACGGTGTGGTGTGACCTGGCGTTCCACGCGCGCGCGC 342  
 QY 497 TTGGCGTCTTTTGTGCGCACATGTTGGTGTGATGATTTCCGAATTTCTACAATGCC 556  
 Db 343 TTGGCGCGCTTCGGGTACCGGTGTGCGCTGCTGTTCTGGCGGTGTTTATGAGGAC 402  
 QY 557 ACGGMAAAGAGCCCGC-----AACTATTGAGGAGACACC 592  
 Db 403 AACCCAAAGCGCCCGAGCCGTTGCGCGTTGCGCCTCGCGGAACCTCTCGCGAGCGG 462  
 QY 593 TCTGTGGCATCTGCTGCTCACCTGGATTCATGTTCCGAAGCTTCGCTGCGATGCTG 652  
 Db 463 TCGCCACCGTCTTCTGCGCTGCTGGGTGCGGTGTTCTGCGCTCTCTGCGCGCGCTG 522  
 QY 653 TCGGTGATGAAAAACAATTCATCCCGGTACATATTTTCATTTGAGGTTTCATGCTGTG 712  
 Db 523 -----GTCTATCCCGCGACGGTCCCGCGGGGTGTTCTGCCGTGATGATCAC 570  
 QY 713 GTATCGCATCGGATGCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATG 772  
 Db 571 GTGTGGCTCTCGACGTGCGCGGTACCGGTGGCGGTGCTGTGTTTGGCAAGCATCCGATG 630  
 QY 773 GCGCGGTGTGTGAGTCCGAAGACTCTTGGGAAGCTTTGCGGCTCCATTGCTTAGGA 832



Db 631 GTCCGCGGATCAGCCCAAGAAATCTCGGAGGCGCTCGCGGCTCGTGTGTGGGT 690  
 QY 833 TCGGTCACTGGTGCATCACTAGTTCACTTCCTGCTCGATCACCATGCTGATGGGTGTG 892  
 Db 691 ATACCGCGGCCACCTCGCGCGAGTTTCTGCGCGGAAGCGCCCGTGGTGGCGCG 750  
 QY 893 ATCTTGGGTGTGECCTAGTTGTGTGCGCCAGCTGGGTGACTTGGTTGAGTTCGCGAGTTC 952  
 Db 751 CTGCTGGGGTGTGTGTGTGTTCACTGSCACCTCGCGGACCTGTGTGAGTTCGCGAGTC 810  
 QY 953 ARAAGGATTTGGGATCAAGGATATGCAACCTTCTTCAAGCCAGCGGATTTGATG 1012  
 Db 811 AAGCGTGAAGTGGGATCAAGGATATGCAACCTTCTTCAAGCCAGCGGATTTGATG 870  
 QY 1013 GACCGTTTGGATGCGATGCTCCCGCGCGGATGCGTGGTGTGATCCTGCA 1064  
 Db 871 GACCGGCTGGAGGTGTGTGTGCTGCGGTGCGCGGTGCGCGCTGAGCGTGTCTGA 922

## RESULT 13

ACA39830  
 ID ACA39830 standard; DNA; 939 BP.

AC ACA39830;

XX 19-JUN-2003 (first entry)

DT Prokaryotic essential gene #21487.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX Mycobacterium leprae.

XX WO200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU35960.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 27700; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing of a gene in an operon required for

CC proliferation or the activity of the gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 939 BP; 144 A; 248 C; 310 G; 237 T; 0 U; 0 Other;

Query Match 12.6%; Score 164; DB 7; Length 939;

Best Local Similarity 52.6%; Pred. No. 4.7e-41;

Matches 440; Conservative 375; Indels 21; Gaps 3;

QY 257 GCGGTGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACTGGGGTCTCTGTTCTT 316

Db 97 GCGGACGTAATCTCCCGCGCGGATCGCGTGGGCTTAAGTATGGTGGTCTTCTGTC 156

QY 317 TTGGGGATTTGCTTAAGCCCATGGGCTTGGTACATCTCTGTTGAGGTTTATGCTGCA 376

Db 157 GCAACGCTGGTGTCTCCGCGAATCTGGTGTCTTGTGCGCTGGCGCATTTTCTGT 216

QY 377 GCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 436

Db 217 GCTAGCCATGAGGTGTGCGCGGCTACGGAAGCTGATATGTGATTCGGCTATCCG 276

QY 437 ATGATCATCGGCGTCAAGCAATCATCTGGCTGTGTCATGGCCATTTGGCAGATGGCAT 496

Db 277 CTGCTCATCGTGGGCGAGTTCACCGTCTGGTAACTTGGCGGTATCGCACCGTTGGTGA 336

QY 497 TTGGGCTCTTTTGGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 550

Db 337 TTGGGCTGTCTTGGTGGCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 396

QY 551 AATGGCACGAAAAGAACGCGCAACTATTGTAGGGACACTCTGTGGGCACTCTCTGTG 610

Db 397 AACACAGACAGCAGCAATCTCGGAGACTTTGGCAGTCCGCGGTGTCAATTAATCTG 456

QY 611 -----CTACCTGGATTCATTTGGGAAGCTTCTGCTGGGATCTGTGCTGATGCAAA 665

Db 457 CGGGACGCGTCCGCCACCGCTTTTGGCGCGGTGGGTCCCGTGTCTTCCCTCGTGGCA 516

QY 666 ACA-----ATTCCATCCCGGTACATATTTCATTTTGAAGTTCATCTGCTGTGTG 715

Db 517 GATTTGCTGGTTTATCCGAGAGTGTGCGGCGCGGTATTCTCCCTGATGATTCGGTG 576

QY 716 ATGCGCATCGGATGTGGCGGGTATATCGCGGTGTGTCTTTTGGATCGCACCCAAATGGCG 775

Db 577 GTAGCTTCGGATGTGGTGGTTACACAGTAGGGGTACTTTTTCGGCAAAACATCCACTGTT 636

QY 776 CGTTTGGTGTGAGTCCGAGCAAGCTTTGGGAAGGCTTTCCGCGCTCCATTGCTTAGGATCG 835

Db 637 CCAAGGATTACCCCGGATTAAGTCTGAGAGGATTTCCCGGTTCGCTGTGCTGCGGAC 696

QY 836 GTCATGTGTGCACTGAGTTTCACTTCTGCTCGATCACCACTGGTGGATGGGTGATC 895

Db 697 ACCGCAACCATCTGACTGTACTCTTCTGCTGGCAAAACACCGTGGGTGGGCTTTG 756

QY 896 TTGGGTGTGCGCTAGTTGTGTCGCGCACGCTGGTGTGCTGTGCTGTGCTGTGCTGTG 955

Db 757 CTCAGTTTGTACTGTGCTCACTGTGACGCTGGGGATTTGGTGGAGTCCCAAGTTAAG 816





219 ATCCCGGCTTATTCAGTTCTGGTGCTAGTAGAAACCGATTTCGGGTGG 142

188 AAGCCCCCATATCAATGCACCCGCTAATGGCTGTAGTAGAAACCGATTTCGGGTGG 142

Search completed: August 17, 2004, 17:09:36  
Job time : 793 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 16:43:55 ; Search time 5204 Seconds  
(without alignments)  
7459.807 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagctcttcttcgaagga.....aggaaaccgattccgggtgg 1300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_estin:\*

3: em\_estin:\*

4: em\_estin:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_estl:\*

11: gb\_estl:\*

12: gb\_estl:\*

13: gb\_estl:\*

14: gb\_estl:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_man:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158.4	12.2	526	28	BH740535
2	56.8	4.4	885	13	BX425603
3	51.6	4.0	712	13	BX416727
4	48.4	3.7	648	14	CB287945

5	48.4	3.7	695	14	CB287953
6	45.8	3.5	627	13	BO560202
7	45.8	3.5	629	13	BY718485
8	45.8	3.5	649	12	BM934705
9	45.8	3.5	773	14	CB289962
10	45.8	3.5	1322	11	BC026505
11	45.8	3.5	1386	29	AY399407
12	45.6	3.5	829	28	BH402821
13	45.2	3.5	435	9	AI844295
14	45	3.5	1201	13	BX381961
15	44.8	3.4	403	12	BJ894397
16	44.8	3.4	614	12	BJ168290
17	44.8	3.4	632	12	BJ602826
18	44.8	3.4	638	12	BJ579676
19	44	3.4	379	28	B67568
20	43.4	3.3	755	12	BJ610177
21	43	3.3	768	14	CF880314
22	43	3.3	823	14	CB907156
23	42.8	3.3	1201	13	BX361080
24	42.6	3.3	812	14	CK235970
25	42.2	3.2	471	12	BJ579141
26	42	3.2	884	29	CNS006U0
27	42	3.2	895	29	CNS0071A
28	41.8	3.2	558	14	CD871811
29	40.8	3.1	1101	29	CNS00L72
30	40.6	3.1	534	13	BX522779
31	40.4	3.1	1201	13	BX355644
32	40.2	3.1	612	13	BU080466
33	40.2	3.1	699	13	BX424825
34	39.8	3.1	698	29	CG897864
35	39.8	3.1	705	29	CG988009
36	39.4	3.0	870	29	AG049346
37	39.4	3.0	1201	13	BX356664
38	39.2	3.0	546	28	BH900070
39	39	3.0	627	12	BM934188
40	39	3.0	849	13	BX462111
41	39	3.0	1101	29	CNS016HG
42	38.8	3.0	534	12	BM129973
43	38.8	3.0	686	13	BM041691
44	38.6	3.0	583	9	AV545973
45	38.6	3.0	703	9	AV783262

#### ALIGNMENTS

RESULT 1  
BH740535  
LOCUS

DEFINITION

BH740535 526 bp DNA linear GSS 22-FEB-2002  
cpbav00114 Corynebacterium pseudotuberculosis Bacterial Artificial  
Chromosomes library Corynebacterium pseudotuberculosis genomic  
clone CB2.4-f10r, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH740535

BH740535.1 GI:18861038

GSS.

Corynebacterium pseudotuberculosis

Corynebacterium pseudotuberculosis

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1 (bases 1 to 526)

Azevedo, V., Fachim, M.S. and Oliveira, S.

Construction and Characterization of Corynebacterium

pseudotuberculosis Bacterial Artificial Chromosomes library

Unpublished (2002)

Contact: Azevedo V

Department of General Biology

Federal University of Minas Gerais

Av. Antonio Carlos, 6627, Campus Pampulha, Belo Horizonte, Minas

Gerais, Brazil

Tel: 00 55 31 3499 2610

Fax: 00 55 31 3499 2610

Email: vasco@mcno.icb.ufmg.br

SEQUENCES WERE BASECALLED WITH PHRED AND VECTOR WAS MASKED WITH

CROSSMATCH. SEQUENCES WERE THEN TRIMMED FROM BOTH ENDS TO REMOVE  
LOW QUALITY BASES AND MASKED VECTOR.  
Seq primer: M13 Reverse  
Class: BAC ends.

# FEATURES

source  
1..526  
/organism="Corynebacterium pseudotuberculosis"  
/mol\_type="genomic DNA"  
/strain="1002"  
/db\_xref="taxon:1719"  
/clone="Cb2.4-f10r"  
/lab\_host="Escherichia coli DH10B"  
/clone\_lib="Corynebacterium pseudotuberculosis Bacterial  
Artificial Chromosomes library"  
/note="Vector: pBelobAC11; Genomic DNA was digested with  
HindIII and randomly cloned into HindIII digested and  
dephosphorylated pBelobAC11."

# ORIGIN

Query Match 12.2%; Score 158.4; DB 28; Length 526;  
Best Local Similarity 66.6%; Pred. No. 2.2e-35; Mismatches 138; Indels 5; Gaps 4;  
Matches 285; Conservative 0;

QY 634 AAGCTTCGTCGATGCTGCTGATGCAAAACAAATCCATCCCGGGTACATATTTTCAT 693  
DB 1 AAGCTTTGAGCTAGCTCTCTCAGTTCAAAATGACGTAGCTCTCTGGGCATATTTGT 60  
QY 694 TTGACGTTTCATGCTGTGTGATCGCATCGATGCGGGGGTATATCGGGGTGTGTT 753  
DB 61 TGTAACTTCATGCTTTGGGTGTTGCTTCTGACGTTGGTGGCTATGTCGCTGTGAT 120  
QY 754 CTTTGGATCGCAACCAATGGCGCGTGTGAGTCCGAGAAAGTCTTGGGAAGCTTTGC 813  
DB 121 GTTTGATCCCAACGATGGCTCTCTGAGTGGTCCAAAGAAATCGTGGGAAGTTTCGC 180  
QY 814 CGGCTCCATTGCTTAGGATCGGTCATGCTGGTGCATCACTAGTTCATCTCTGCTCGATCA 873  
DB 181 GGGTTTCAGTTATTTTGGAATGACTGTCGGCGTATTGAGTGTGATTTCTCTCGGATTA 240  
QY 874 CCACGTGGTGGTGTGATCTGGTGTGCTGAGTGTGCGCGACGTTGGGTGA 933  
DB 241 CACCTGGTGGTGGGCTTAATCTAGGTTGGATTTGGCTATCTTACCTTAGCGGA 300  
QY 934 CTTGTGTGA-GTCGCAGTTCAAAACCGATTTGGGCATCAAGGATATGTCGAACCTTC-TT 991  
DB 301 TTTAGTGGANGTCAAGTTCAGCGAGAACTAGGTATTAGGACATGTCGCTATGCTTT 360  
QY 992 CCAGGCCACGGCGG--ATTGATGACCGTTTGGATGGCATGCT-CCCGGCCGCGATGGT 1048  
DB 361 CCCGCCACGGTGGGTCTTATTGGATAGCTCCGACGGAATGTGCCCTTCTGCCATGTT 420  
QY 1049 ACGTGGTT 1056  
DB 421 ACGTGGGT 428

RESULT 2  
BX425603  
LOCUS  
DEFINITION  
BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone  
CLOBB0222A07 3-PRIME, mRNA sequence.

ACCESSION  
BX425603  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 885)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paraday Avenue Genoscope sequence ID : CLOBB0222A07FPI.

# FEATURES

source  
1..885  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CLOBB0222A07"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

# ORIGIN

Query Match 4.4%; Score 56.8; DB 13; Length 885;  
Best Local Similarity 3.3%; Pred. No. 5.2e-05; Mismatches 312; Indels 0; Gaps 0;  
Matches 16; Conservative 161;

QY 338 TGGGGTTGTACATCCTCGTTCAGGTTTATGGCTGCGACCAACATGGAGTTCGTAGC 397  
DB 339 TNNKKKKNTTKTKTKKKKKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 398  
QY 398 AGACTTAAAGAGGGCGGTATCATTTGCCACTGCGCGATATGATCATCGCGGTGAGCA 457  
DB 399 NNANNNANNNANNN 458  
QY 458 ATCATCTGCTGCTGATGGCCATTTCGACGATGGGCATTTTGGCGCTCTTTTGGCCACT 517  
DB 459 KKK 518  
QY 518 GTGTGCTGCTGATGTTTCGGAATTTTCAATATGGCACGGAAGAAAGAGCCGCGAAC 577  
DB 519 KKK 578  
QY 578 TATTGAGGACACCTCTCTGCGGCATCTCGTCTCACCTGATTCATTCATTTTTCGGAAGC 637  
DB 579 NNN 638  
QY 638 TTGCTGCGATGCTGCTGATGCAAAACAAATTCATCCCGGGTACATATTTTCATTTTG 697  
DB 639 KNN 698  
QY 698 ACCTTCATCTGCTGATGATCGCATGCGATGTTGGGGGGTATATCGCGGTGTGTTT 757  
DB 699 NNN 758  
QY 758 GGATCGCACCAATGCGCGCTTGGTGGTCCGAGAGAGTCTTGGGAAGGCTTTGCGCGC 817  
DB 759 KKK 818  
QY 818 TCCATTGTC 826  
DB 819 KNNNNNNNN 827

RESULT 3  
BX416727  
LOCUS  
DEFINITION  
BX416727 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone  
CS0DA011V114 5-PRIME, mRNA sequence.

ACCESSION  
BX416727  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens





[REDACTED]





bacteria (Life Technologies) to generate the NIH BMAP M.S3.2 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

## ORIGIN

Query Match 3.5%; Score 45.8; DB 12; Length 649;  
Best Local Similarity 53.3%; Pred. No. 0.08;  
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 915 TGTGGCCACGTTGGGTGACTTGGTTGAGTCGCGAGTTCAAAACGCGATTGGGCATCAAGG 974  
DB 200 TGATGGCCCATTTGGAGGCTCTTTCCAGTGATTCAAAAGAGCTTCAAAATCAAGG 259  
QY 975 ATATGTCGAACCTTCTTCCAGGCCACGCGGAGTGTGATGGACCGTTTGGATGGCATGC--- 1031  
DB 260 ATTTGCAACACAGATTCTCTGACATGCGGGATATGGACAGGTTTGACTGTCAGTATT 319  
QY 1032 TCCGGCCGGGATGGTACGTTGTTGATCTGATGTGATGATCAGCAGCTCGTATCCGTGCT 1091  
DB 320 TGATGGCAACGTTTCGTGCACGTTGATCATCACCAGCTTCATCAGGGGTCCGAATCCGAGCA 379  
QY 1092 AAAGCTTGGGCCACCTTTAAGTTCAAAAACCTTGAAAGGCGCTGA 1136  
DB 380 AAGTGTCCAGCAGCTGTAGTGTCTTCAACGACAGCAGTTAA 424

RESULT 9  
CB289962 779 bp mRNA linear EST 09-JUL-2003  
LOCUS UI-M-POO-cdv-n-05-0-UI.r1 NIH\_BMAP\_F00 Mus musculus CDNA clone  
DEFINITION IMAGE:6835110 5', mRNA sequence.

ACCESSION CB289962  
VERSION CB289962.1 GI:28611377

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 779)

NIH-MGC Project URL: <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seg primer: pYX-5.

Location/Qualifiers

1. 779

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:6835110"

/tissue\_type="whole brain"

/dev\_stage="embryo 12.5dpc"

/lab\_host="DH10B (T1 phage resistant)"

/clone\_lib="NIH\_BMAP\_F00"

/note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I;

Site\_2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TCAGAGAGCC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Query Match 3.5%; Score 45.8; DB 14; Length 779;  
Best Local Similarity 53.3%; Pred. No. 0.09;  
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 915 TGTGGCCACGTTGGGTGACTTGGTTGAGTCGCGAGTTCAAAACGCGATTGGGCATCAAGG 974  
DB 70 TGATGGCCCATTTGGAGGCTCTTTGCCAGTGGGATTCAAAGAGCTTCAAATCAAGG 129  
QY 975 ATATGTCGAACCTTCTTCCAGGCCACGCGGAGTGTGATGGACCGTTTGGATGGCATGC--- 1031  
DB 130 ATTTGCAACACAGATTCTCTGACATGCGGGATATGGACAGGTTTGACTGTCAGTATT 189  
QY 1032 TCCGGCCGGGATGGTACGTTGTTGATCCTGAGTGTGATCAGCAGCTCGTATCCGTGCT 1091  
DB 190 TGATGGCAACGTTTCGTGCACGTTACATCACCAGCTTCATCAGGGGTCCGAATCCGAGCA 249  
QY 1092 AAAGCTTGGGCCACGCTTTAAGTTCAAAAACCTTGAAAGGCGCTGA 1136  
DB 250 AAGTGTCCAGCAGCTGTAGTGTCTTCAACGACAGCAGTTAA 294

## RESULT 10

BC026505

Mus musculus, clone IMAGE:4507928, mRNA.

ACCESSION BC026505

VERSION BC026505.1 GI:20071490

KEYWORDS HTC.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1322)

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 44 Row: i Column: 9

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis

This clone has the following problem: frame shifted.

Location/Qualifiers

1. 1322

/organism="Mus musculus"

/mol\_type="mRNA"

## FEATURES

source

```

/db_xref="taxon:10090"
/clone="IMAGE:4507928"
/tissue_type="Eye, retina, mouse strain C57Bl\6"
/clone_lib="NIH MGC_94"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"

ORIGIN
Query Match      3.5%; Score 45.8; DB 11; Length 1322;
Best Local Similarity 53.3%; Pred. No. 0.12;
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 915 TGTGGCCACGTTGGGTGAGTCTTCCAGGCCCGGGGATTCAGTCCAGTTCACAAAGCGATTTGGGCAATCAAGG 974
Db 504 TGATTGGCCCATTTGAGGCTTCTTTGCCAGTGGATTCAAAAGAGCTTTCAAATCAAGG 563
QY 975 ATATGTGGAACCTTCTTCCAGGCCCGGGGATTCAGTCCAGTTCACAAAGCGATTTGGGCAATCAAGG 1031
Db 564 ATTTTGAACACAGATTCCTGGACATGGCGGATATGACAGAGTTTGACTGTCAATATT 623
QY 1032 TCCCGCGCGGATGTGAGTGTGATCTGAGTGTGATCAGCAGCTCGTATCCGTCGT 1091
Db 624 TGATGGCAACGTTGCTGACGTGTACATCACCAGCTTCATCAGGGTCCGAATCCAGCA 683
QY 1092 AAAGCTTGGCCAGCTTTAAAGTTCAAAAACCTTGAAGCGGCTGA 1136
Db 684 AAGTCTCCAGCAGCTGTAGTCTTCAACAGACAGCAGTTAA 728

RESULT 11
AY399407 1386 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CDS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION AY399407
ACCESSION AY399407
VERSION AY399407.1 GI:39755396
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1386)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14671302
REFERENCE 2 (bases 1 to 1386)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1386
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1386
/gene="CDS1"
/locus_tag="HMC0213"

ORIGIN
Query Match      3.5%; Score 45.8; DB 29; Length 1386;
Best Local Similarity 53.3%; Pred. No. 0.13;
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

```

```

QY 915 TGTGGCCACGTTGGGTGAGTCTTCCAGGCCCGGGGATTCAGTCCAGTTCACAAAGCGATTTGGGCAATCAAGG 974
Db 1094 TGATTGGCCCATTTGAGGCTTCTTTGCCAGTGGATTCAAAAGAGCTTTCAAATCAAGG 1153
QY 975 ATATGTGGAACCTTCTTCCAGGCCCGGGGATTCAGTCCAGTTCACAAAGCGATTTGGGCAATCAAGG 1031
Db 1154 ATTTTGAACACAGATTCCTGGACATGGCGGATATGACAGAGTTTGACTGTCAATATT 1213
QY 1032 TCCCGCGCGGATGTGAGTGTGATCTGAGTGTGATCAGCAGCTCGTATCCGTCGT 1091
Db 1214 TGATGGCAACGTTGCTGACGTGTACATCACCAGCTTCATCAGGGTCCGAATCCAGCA 1273
QY 1092 AAAGCTTGGCCAGCTTTAAAGTTCAAAAACCTTGAAGCGGCTGA 1136
Db 1274 AAGTCTCCAGCAGCTGTAGTCTTCAACAGACAGCAGTTAA 1318

RESULT 12
BH402821/c 829 bp DNA linear GSS 11-DEC-2001
LOCUS AG-ND-178A15, TR ND-TAM Anopheles gambiae genomic clone
DEFINITION AG-ND-178A15, genomic survey sequence.
ACCESSION BH402821
VERSION BH402821.1 GI:17349037
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
1 (bases 1 to 829)
Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
Mol. Genet. Genomics 268 (6), 720-728 (2003)
JOURNAL PUBMED 12655398
MEDLINE 22542063
COMMENT Other GSSs: AG-ND-178A15.TF
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PST strain
(DNA). All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University, College Station, Texas 77843-2123, USA using a HindIII
partial digest.
Seq primer: M13 Rev
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..829
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PST"
/db_xref="taxon:7165"
/clone="AG-ND-178A15"
/clone_lib="ND-TAM"
/notes="Vector: pSCBAC1; Site_1: HindIII"

ORIGIN
Query Match      3.5%; Score 45.6; DB 28; Length 829;
Best Local Similarity 62.1%; Pred. No. 0.11;

```

```

Matches 72; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 926 TTGGTACTGTGTTGAGTCGAGTTCAAACGGGATTTGGGCATCAGGATATGTGAAC 985
Db 795 TTGGGAGATTGGTGGAGTACAGCTCAAAGAACCTTTCAATGTAAGAAGATTCGGGGAA 736
QY 986 CTTCTTCAGCCACGGCGGATTCATGACCGCTTTGATGGCATGCTCCCGGCGC 1041
Db 735 ATAATCCGGGACGAGGTGATTCGACAGGCTGGATAGCTTATACCTTTGGC 680

RESULT 13
A1844295 435 bp mRNA linear EST 15-JUL-1999
LOCUS UI-M-AL1-ahm-d-04-0-UI.s1 NIH BMAP_MCO.N Mus musculus cDNA clone
DEFINITION UI-M-AL1-ahm-d-04-0-UI 3', mRNA sequence.
ACCESSION A1844295
VERSION A1844295.1 GI:5488201
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 435)
COMMENT Bernaldo, M.F., Lennon, G. and Soares, M.B.
Normalizaton and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE 889548
PUBMED Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mst@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. When NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
Seq primer: M13 forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..435
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="UI-M-AL1-ahm-d-04-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_MCO_N"
/note="vector: pT7TD-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_MCO_N library is a normalized library constructed
from mouse cortex. The tag is a string of 5 nucleotides
present between the Not I site and the oligo-dt track.
The library was constructed as described by Bernaldo,
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Ms. Annie Novakovich, Zivic-Miller
Laboratories.
TAG_TISSUE=prefrontal-cortex
TAG_LIB=NIH_BMAP_MCO_N
TAG_SEQ=GCTCA"

ORIGIN
Query Match 3.5%; Score 45.2; DB 9; Length 435;
Best Local Similarity 53.7%; Pred. No. 0.095;
Matches 117; Conservative 0; Mismatches 98; Indels 3; Gaps 1;


```

```

QY 922 CACGTTGGTGGTACTTGGTTGAGTCGAGTTCAAACGGATTTGGGCATCAAGGATATGTC 981
Db 9 CTCATTGGAGGCTTCTTTGGCCAGTGGATTCAAAGAGCTTTCAAATCAAGGATTTGTC 68
QY 982 GAACCTTCTTCAGGCCACCGCGGATTCATGACCGTTGGATGGCATGC---TCCCGGC 1038
Db 69 AAACACGATTCTCGACATGGCGGATAATGACACAGGTTTGACTGTGATTTTCATGTCG 128
QY 1039 CCGGATGGTGACGTGGTTGATCTCAGTGTGATCAGCAGCTGATCGTTCGTGAAGCTT 1098
Db 129 AACGTTGGTGACGTGTATACATCACCAGCTTCAACAGGAGTCCGAATCCAGCAAAAGTCT 188
QY 1099 GGGCCAGCTTTTAAGTTCAAAAACCTTGAAGGCGCTGA 1136
Db 189 CCAGCAGCTGTTAGTGTCTTCAACACGAGACAGCAGTTAA 226

RESULT 14
EX381961 1201 bp mRNA linear EST 08-MAY-2003
LOCUS EX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI072YF05 3-PRIME, mRNA sequence.
ACCESSION EX381961
VERSION EX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1201)
COMMENT Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CSODI072CC03NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 3.5%; Score 45; DB 13; Length 1201;
Best Local Similarity 5.3%; Pred. No. 0.2;
Matches 32; Conservative 223; Mismatches 373; Indels 2; Gaps 1;
QY 64 TAAGATTAACGTGTTTCAACGTGATTTGGCGGGGAAACGACGCTTTCTTTTGTGCTGCA 123
Db 1077 KVVVGMNNKKKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMM 1018
QY 124 AGAGTGTGTTGGAGAAATTTTTCGAAATGCTGGCCACCATCAACAGTGCATGTTAGAA 183
Db 1017 KTGKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 958
QY 184 ACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCCCAA 243
Db 957 KMKMKMGVNNKTKMMMKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNK 898
QY 244 ACCCAAAAATAATACGGGTCGAGATCTCAAACTGCCATGCCATCGGATCGGATGGG 303
Db 1077 KVVVGMNNKKKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMM 1018

```



Query Match	14.8%;	Score 191.8;	DB 3;	Length 4403765;
Best Local Similarity	54.4%;	Pred. No. 9.6e-52;		
Matches 456;	Conservative 0;	Mismatches 347;	Indels 36;	Gaps 2;
Qy	257	GCGGGTGAGATCTCAAAGCTGCATCTGTGTGGGATCGGACTGGGGGTCTCTGGTTCCT	316	
Db	3185640	GC GGCCGTCATTTGCGCGCAGGATCGTGTGGGCTTTCCATAGCCCTAGTCCTCATC	3185581	
Qy	317	TTGGGATGTGCTTAAGCCATGGGGTTGGTACATCTCTGTTGCAGGTTTATGGCTGCA	376	
Db	3185580	GC GGTCGTGCTGTTCGTCGCGGTTTGGGTTGCCATCGTGGCCGTGCGCACCTTGGTC	3185521	
Qy	377	GCAACATGGGAAGTTCGTATAGCAGACTTAAAGAAGCGGGCTATCATTTGGCACTTGC CGAAT	436	
Db	3185520	GCTTACCCTAGGTGTGCGGAGTTTCGGGAAGCGGGCTATCTCATCCGGTTATCCCG	3185461	
Qy	437	ATGATCATCGGGTCAGCAATCATCTGGCTGTATGCGCCATTTGGCAGATGGGCAT	496	
Db	3185460	TTGCTGATTTGGCGGGCAGCCCGCGTGTGGCTGACCTGGCCGTTTCGGCGGTCGCGCGCA	3185401	
Qy	497	TTGGCGTCTTTTGTGCCACTGTGTGGTGCCTGAGTATTTCCGAAATTTCTACAATGCC	556	
Db	3185400	TTGGCGGGCTTTGGTGGCATGTGCTGTGTCTGCATGATTTGGCGACTGTTTCATGCAGGAC	3185341	



QY 690 TCATTTTGACGTTTCATGCTGTGTGATCGCATCGATGTGGCGGGTATATCGCGGGTG 749  
Db 1097 TGATCATCCGCTGATGGTGTGGTCTGGGCGCGGATATCGGCGCTACTTCTCCGGCA 1156  
QY 750 TGTCTTTGGATCGACCCAAATGGCGCGCTTGGTGGTCCGGAAGTCTTTGGGAAGCT 809  
Db 1157 AGGCTTTCGGCAAGCGCAAGCTTGGCGCGCGGTTCAGCCCCCGCAAGAGCTGGGAAGGCG 1216  
QY 810 TTGCGGCTCCATTTCTTAGGATCGG---TCACTGGTGTCACTCAGTGTTCACCTTCCTGC 866  
Db 1217 TTACGCGCGCTGGCGGCACGCTGGCGATCACCTGGCGGTTCACCGGGCT 1276  
QY 867 TCGATCACCACTGGTGATGGGTGTGATCTTGGGTGTGCCCTAGTGTGTGCGGCAAGT 926  
Db 1277 GGTCCCTCGCGCGCTGTCTTGGCGCGCTGTGGCGCGCGCTGTGGTGTGTCTCGA 1336  
QY 927 TG---CGTGACTTGGTGTGATCGCACTTCAAGCGGATTTGGGCATCAAGGATATGTGCA 983  
Db 1337 TGTGCGGACCTTACCGAAGCATGTTCAAGCGGCAATCCGGAATCAAGGACGAGCA 1396  
QY 984 ACCTTCTTCCAGGCCACGGCGGATTTGATGACCGGTTTGGATGGCATGCTCCCGGCGGCA 1043  
Db 1397 ACTGTGTCGCGCCACGCTGGCGGTGTGATCGCATCGACAGCCTGACGCGAGCATCC 1456  
QY 1044 TGGTG 1048  
Db 1457 CGGTG 1461

RESULT 4  
US-09-252-991A-4542  
; Sequence 4542, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4542  
; LENGTH: 1854  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4542

Query Match 4.9%; Score 64.2; DB 4; Length 1854;  
Best Local Similarity 52.3%; Pred. No. 6.8e-11;  
Matches 191; Conservative 0; Mismatches 168; Indels 6; Gaps 2;  
QY 690 TCATTTTGACGTTTCATGCTGTGTGATCGCATCGGATGTGGCGGGTATATCGCGGGTG 749  
Db 261 TGATCATCGCGTGTGCTGTGTGGCGCGGATATCGGCGCTACTTCTCCGGCA 320  
QY 750 TGTCTTTGGATCGACCCAAATGGCGCGCTTGGTGGTCCGGAAGTCTTTGGGAAGGCT 809  
Db 321 AGGCTTTCGGCAAGCGCAAGCTTGGCGCGCGGTTCAGCCCCCGCAAGAGCTGGGAAGGCG 380  
QY 810 TTGCGGCTCCATTTCTTAGGATCGG---TCACTGGTGTCACTCAGTGTTCACCTTCCTGC 866  
Db 381 TTACGCGGCTGTGGCGGCGAGCTGGGATACCCCTGGCGGTTCGGCTCTACCGGGCT 440  
QY 867 TCGATCACCACTGTGTGATGGGTGTGATCTTGGGTGTGCCCTAGTGTGTGCGGCAAGT 926  
Db 441 GGTCCCTCGGCGCGCTGTCTTGGCGGCTGTCTGCGCGGCGCTGGTGTGTCTCTCGA 500  
QY 927 TG---GGTGACTTGGTGTGATCGGCAAGTTCAAACCGGATTTGGGCAATCAAGGATATGTGCA 983

Db 501 TCGTCGGCGACCTTACCGAAGCATGTTCAAGCGGCAATCGGAATCAAGGACAGCAGCA 560  
QY 984 ACCTTCTTCCAGGCCACGGCGGATTTGATGACCGTTTGGATGGCATGCTCCCGCGCGCA 1043  
Db 561 ACCTGTCTCCCGGCGCACGCTGGCGTGTGATGCGATCGACAGCTTACGCGCAGCCATCC 620  
QY 1044 TGGTG 1048  
Db 621 CGGTG 625

RESULT 5  
US-09-252-991A-14536  
; Sequence 14536, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 14536  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14536

Query Match 4.5%; Score 58.8; DB 4; Length 1041;  
Best Local Similarity 50.9%; Pred. No. 3.1e-09;  
Matches 166; Conservative 0; Mismatches 157; Indels 3; Gaps 1;  
QY 722 TCGGATGTGGCGGGTATATCGCGGTGTGTTTGGATCGCACCCCAATGGCGCGGTG 781  
Db 692 TCCGACGTGATCGATACATCTCGGCAAGCTGTTCGCAAGCGCAAGATCGTCCCAAC 741  
QY 782 GTGATCGGAAGATCTTGGAGGCTTTCGGGCTCCATGTTCTTAGGATCGGTCACT 841  
Db 742 CTGTGCGGCTCGAAGACCGTTCGAAGGCTTCGTGCGCGCATCGCCCTGCCACCCCATC 801  
QY 842 GGTGCACTCAGTGTTCACCTTCTGCTCGATCACCACTGTTGGATGGGTGATGTTGGGT 901  
Db 802 GCGGCTGCTGTGTGATCACCCGTTCAACCTTGGCAGGCGGCTTGTATCGCGGTG 861  
QY 902 TGTGCCCTAGTGTGTGCGCCACGTTGGGTGACTTGGTGTGATGTCAGTTCGCGGAT 961  
Db 862 CTGATCAACCTTCTGGGCTTC---TTCGCGGCTTGTGATGTCGCGCATCAAGCGCGAC 918  
QY 962 TTGGGCAATCAAGATATGTCGAACCTTCTCCAGGCCACGGCGGATTTGATGACCGTTG 1021  
Db 919 CBTGGGTCAAGGACTTGGGGGCAATGATCGAAGTTCACGTTGGCATGCTCCACCGCTG 978  
QY 1022 GATGGCATCTCCCGCGCCGCGATGTT 1047  
Db 979 GACTCGGTGTGCTTCGCGCGCGGAT 1004

RESULT 6  
US-09-252-991A-14831/c  
; Sequence 14831, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A



```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14831
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14831

Query Match
Best Local Similarity 4.5%; Score 58.8; DB 4; Length 1209;
Matches 166; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 722 TCGGATGTGGCGGGTATATCGCGGGTGTGTTCTTTGGATCGCACCAATGGCGCCGTTG 781
DB 480 TCCGACGTGATGACGTACATCTCGGCAAGCTGTTGGCAAGCGCAAGATCGCTCCCAAC 421
QY 782 GTGAGTCCGAAGAAGTCTTGGGAAGCTTTCGGCTCCATGTTCTTAGATCGGTCACT 841
DB 420 CTGTGCGCGTGCNAGACCGTGCNAGAGGCTTCTGCGCGGCGATCGCCCTGGCCACCGCCATC 361
QY 842 GGTGCACTCAGTGTCTCTCTCTCGATCACCACTGGTGGATGGTGTGATCTTGGGT 901
DB 360 GCGGCTCGTGTGGTGAATCACCCGTTCAACCTTGGCAGGCGGTTGATCGCGCTG 301
QY 902 TGTGCCCTAGTGTGTGGCGCACGTTGGGTGACTTGGTGGATGGTGGATGCAACCGCAT 961
DB 300 CTGATCAACCTGTGGGTTCTTCCAGGCGGCGTGTGATGTCGGCGCATCAAGCGGCAC 244
QY 962 TTGGGATCAAGGATGTGAACCTTCTTCCAGGCGCACGCGGATGATGGACCGTTG 1021
DB 243 CGTGGGTCAAGGACTGGGGGACATGATCAAGGTCAAGGTGCGATGCTCGACCGCGT 184
QY 1022 GATGCGATGTCGCGCGCGGATGTT 1047
DB 183 GACTCGGTGTCTTCCGCGCGCGAT 158

RESULT 7
US-09-489-039A-5219
; Sequence 5219, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5219
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5219

Query Match
Best Local Similarity 4.5%; Score 58.2; DB 4; Length 870;
Matches 187; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 702 TCATGCTGTGTGATCGCATCGGATGCGCGGGTATATCGCGGTGTGTTCTTTGGAT 761
DB 482 TCATGATCTCTGTGTGGGCGGACATCTTGGCGCTATATGTTGGTAAATGTTTGGCA 541
QY 762 CGCACCAATGGCGCGGTGTGGTGGTCCGAAGAAGTCTTGGGAAGC---TTTGGCGGT 818
DB 542 AGCACAAAGTGGCGGCGGAGGTCTACCCGGTAAACCTTGGCAGGCGTTTGTGGCGTT 601

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14831
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14831

Query Match
Best Local Similarity 4.5%; Score 58.8; DB 4; Length 1209;
Matches 166; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 722 TCGGATGTGGCGGGTATATCGCGGGTGTGTTCTTTGGATCGCACCAATGGCGCCGTTG 781
DB 480 TCCGACGTGATGACGTACATCTCGGCAAGCTGTTGGCAAGCGCAAGATCGCTCCCAAC 421
QY 782 GTGAGTCCGAAGAAGTCTTGGGAAGCTTTCGGCTCCATGTTCTTAGATCGGTCACT 841
DB 420 CTGTGCGCGTGCNAGACCGTGCNAGAGGCTTCTGCGCGGCGATCGCCCTGGCCACCGCCATC 361
QY 842 GGTGCACTCAGTGTCTCTCTCTCGATCACCACTGGTGGATGGTGTGATCTTGGGT 901
DB 360 GCGGCTCGTGTGGTGAATCACCCGTTCAACCTTGGCAGGCGGTTGATCGCGCTG 301
QY 902 TGTGCCCTAGTGTGTGGCGCACGTTGGGTGACTTGGTGGATGGTGGATGCAACCGCAT 961
DB 300 CTGATCAACCTGTGGGTTCTTCCAGGCGGCGTGTGATGTCGGCGCATCAAGCGGCAC 244
QY 962 TTGGGATCAAGGATGTGAACCTTCTTCCAGGCGCACGCGGATGATGGACCGTTG 1021
DB 243 CGTGGGTCAAGGACTGGGGGACATGATCAAGGTCAAGGTGCGATGCTCGACCGCGT 184
QY 1022 GATGCGATGTCGCGCGCGGATGTT 1047
DB 183 GACTCGGTGTCTTCCGCGCGCGAT 158

RESULT 8
US-09-221-017B-479/c
; Sequence 479, Application US/09221017B
; Patent No. 6444739
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PPI182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PPI546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 479:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

```



ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE:  
ORGANISM: PORPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...7060  
US-09-221-017B-479

Query Match 4.2%; Score 54.4; DB 4; Length 7060;  
Best Local Similarity 50.6%; Pred. No. 3.3e-07;  
Matches 158; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

QY 725 GATGCGCGGGTATATCGCGGGTGTCTTTTGGATCGACCAATGCGCGCGTGGTG 784  
DB 3245 GATACAGAGCTTTATGTCAGGTTCCTCTTGGCAAGCATACACTCTTCAAGGTCA 3186  
QY 785 AGTCGAGAGCTTTGGGAAGCTTTCCGGCTCCATTGCTTAGATCGGTCACTGT 844  
DB 3185 TCACCCAGAAAACATGGGAGGGCTTTATCGGAGGGCTGCTGCTACTGTGGCAGGTGCA 3126  
QY 845 GCACTAGTGTTCCTCTCTGATCACCACACTGTGTGGTGTGTGATCTTGGGTGT 904  
DB 3125 TTGCTGCAGGCCACTATTCGGGTTCGCCAAGCATCCGATGGAGCTGTCTTTTGTCT 3066  
QY 905 GCCTAGTGTGTGC-----GCCAGTGGGTGACTTGTGTCGCGAGTTCAAACGGAT 961  
DB 3065 CTACTGTGTAACGGCCATGGGCCACTTGGGGCGGATCTGTATGAGTCCAACTCTCAAACGCAAT 3006  
QY 962 TTGGGCTCAAGGATATGTGCACTCTTCTCCAGGCCAGCGGATTTGATGACCGTTTG 1021  
DB 3005 GCAGGAGTGAAGATTCCGGCCATATCATTTCCGGACATGAGGTATATCTGGACCGGTG 2946  
QY 1022 GATGGCATGCTC 1033  
DB 2945 GACAGTGTCTCTC 2934

RESULT 9  
US-09-540-236-198  
; Sequence 198, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 198  
; TYPE: DNA  
; ORGANISM: M. catarrhalis  
US-09-540-236-198

Query Match 4.1%; Score 53.2; DB 4; Length 891;  
Best Local Similarity 52.2%; Pred. No. 2.2e-07;  
Matches 168; Conservative 0; Mismatches 148; Indels 6; Gaps 2;

QY 723 CGGATGTGGCGGGTATATCGCGGGTGTCTTTTGGATCGACCAATGCGCGCGTGG 782  
DB 467 CTGATAGTGGGGCTTATTTTATTTGGCAAGAAATTTGCAAAACGAAATAATGGCAACCAATG 526  
QY 783 TGAGTCCGAGAGTCTTGGGAGGCTT---TGGCGGCTCCATTGCTTAGGATCGGTCA 839  
DB 527 TATCACCACAAATAAAGATATCGAAGGCTTATCGGGTGAATGATGGTCTCTGGGGTGGTCA 586  
QY 840 CTGGTGCACCTCAGTGTTCACCTCTCTGCTCGATCACCACCTGCTG---ATGGGTGTGATCT 896  
DB 587 CTGGGAGTGGTGTATTTATTTGACGTGTCTGGTGTGATTTGTTGATTTTGGATTT 646  
QY 897 TGGGTGTGCCCTAGTGTGTGCGCCACGTGGGTGAGTGTGAGTGCAGTTCACCAAC 956

DB 647 TGTCTTGGCTRACTGTACTTGCAGTATATCTGGCGGACTTATTTGAATCTATGCTAAAC 706  
QY 957 GCGATTTGGGATCAAGGATATGTGAACCTTCTTCCAGGCCACGCGGATGATGGACC 1016  
DB 707 GCGGTGCAGGTATTAAGATTTCTGGGACAAATCTACAGGTCAATGCGGTATTTTGGATC 766  
QY 1017 GTTTGATGGCATGCTCCCGGC 1038  
DB 767 GAATTGATTCGCTACTTTTCAGC 788

RESULT 10  
US-09-596-002-37  
; Sequence 37, Application US/09596002  
; Patent No. 6632636  
; GENERAL INFORMATION:  
; APPLICANT: Lagace, Robert, E.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Berg, Kim, L.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
; FILE REFERENCE: PM-0008-4 US  
; CURRENT APPLICATION NUMBER: US/09/596,002  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: 60/140,121  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PERL Program  
; SEQ ID NO 37  
; LENGTH: 99629  
; TYPE: DNA  
; ORGANISM: M. catarrhalis  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 78467  
; OTHER INFORMATION: a or g or c or t, unknown, or other  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte template ID No. 6632636 37  
; PUBLICATION INFORMATION:  
US-09-596-002-37

Query Match 4.1%; Score 53.2; DB 4; Length 99629;  
Best Local Similarity 52.2%; Pred. No. 4.8e-06;  
Matches 168; Conservative 0; Mismatches 148; Indels 6; Gaps 2;

QY 723 CGGATGTGGCGGGTATATCGCGGGTGTCTTTTGGATCGACCAATGCGCGCGTGG 782  
DB 62838 CTGATAGTGGGGCTTATTTTATTTGGCAGAAATTTGGCAAAACGAAATAATGGCAACCAATG 62897  
QY 783 TGAGTCCGAGAGTCTTGGGAGGCTT---TGGCGGCTCCATTGCTTAGGATCGGTCA 839  
DB 62898 TATCACCACAAATAAAGATATCGAAGGCTTATCGGGTGGATGATGCTCTCTGGGGTGGTCA 62957  
QY 840 CTGGTGCACCTCAGTGTTCACCTTCTGCTCGATCACCACCTGGTGG---ATGGGTGTGATCT 896  
DB 62958 CTGTGCGAGTGTGTTATTTATTTGACGCTGTCTGGTATGTCATTTGTTTGTATTT 63017  
QY 897 TGGGTGTGCCCTAGTGTGTGCGCCACGTGGGTGAGTCTGTTGAGTGCAGGTTCACCAAC 956  
DB 63018 TGTCTTGGCTAACTGCTACTTCTGCCAGTATATCTGGCGGACTTATTTGAATCTATGCTAAAC 63077  
QY 957 GCGATTTGGGATCAAGGATATGTGAACCTTCTTCCAGGCCACGCGGATTTGATGGACC 1016  
DB 63078 GCGGTGCAGGTATTAAGATTTCTGGGACAAATTTACAGGTCAATGCGCGGTATTTTGGATC 63137  
QY 1017 GTTTGATGGCATGCTCCCGGC 1038  
DB 63138 GAATTGATTCGCTACTTTTCAGC 63159

RESULT 11  
US-09-543-681A-2323  
; Sequence 2323, Application US/09543681A  
; Patent No. 6605709

GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543.681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 2323  
LENGTH: 879  
TYPE: DNA  
ORGANISM: Proteus mirabilis  
US-09-543-681A-2323

Query Match 3.9%; Score 50.8; DB 4; Length 879;  
Best Local Similarity 47.9%; Pred. No. 1.4e-06;  
Matches 179; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 668 AATTCATCCCGGTACATATTTTCATTTTGACGTTTCATGCTGTGTGATCGATCGGAT 727  
DB 451 AACTCATATACAGGGCTTGTGTGTACTGTATGTCATGTTATGTTTGGGCTGGGAC 510  
QY 728 GTGGCGGGTATATCGGGGTGTTCTTTGGATCGACCAATGCGCGGTTGGTGGT 787  
DB 511 TCAGGGGCTTATGCAATTTGGCGGTTTAAATGGGTAACATAAAATGGCAGCAAGTATCA 570  
QY 788 CCGAAGAGTTTCGGGAGGCTTTGCCGGCTCCATGCTTAGTATCGGTCACGTGGTCA 847  
DB 571 CCGGTAACCTTGGAGGCTGTGGTGGCGGGTAACTACCTGACGGTGTGTTTCATGG 630  
QY 848 CTCAGTGTTCATTCCTGCTCGATCACCACCTGFGATGCTGTGTGATCTTGGTGTGGCC 907  
DB 631 CTATTTACGGCTTTGCGACCAATTAATGAAGTGCCTAACCATCTATTACTGATTTCCGGC 690  
QY 908 CTAGTTGTGTCG---CCAGTGGGTGACTTGTGTGAGTCGCGAGTTCAACCGGATTG 964  
DB 691 ATCGTGTGATTGTTCCGTATTCGGCGATCTGCGTGAAGATGTTTAAACGTGTTTC 750  
QY 965 GGCATCAAGGATATGTCAACCTTTCTCCAGGCCACGGCGGATGATGGACCGTTGGAT 1024  
DB 751 GGCATTAAGACAGTAGTCAGCTTATCCCTGGCGATGCTGCTGTAGATCGTATGAT 810  
QY 1025 GGCATGCTCCCGC 1038  
DB 811 AGTTTGACCGCGC 824

RESULT 12  
US-09-328-352-3131  
Sequence 3131, Application US/09/28352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03EA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 3131  
LENGTH: 858  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-3131

Query Match 3.5%; Score 45.4; DB 4; Length 858;  
Best Local Similarity 56.3%; Pred. No. 9.1e-05;  
Matches 85; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 891 TGATCTGGGTGGCCCTAGTCTGTGGCCAGCTTGGGTGACTGGTGGATCGACGT 950

DB 653 TGATTTTGTCTTTAACTACTACGTGTTCCGTTTCAGTATTAGGTGATTTATTGAATCAATGA 712  
QY 951 TCACACCGGATTTGGCATCAAGATATGTGCAACCTTCTTCCAGGCCACGCGGATGA 1010  
DB 713 TCACACCGCGCTGCTGGTATTAAAGACTCTGCGCGTGTGTTTCCAGGTCATGGTGGTAT 772  
QY 1011 TGGACCGTTTGGATGGCATGCTCCCGGCGC 1041  
DB 773 TGGACCGTATTGATCTTTACTCCGACGAGC 803

RESULT 13  
US-09-252-991A-4311/c  
Sequence 4311, Application US/09/252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 4311  
LENGTH: 1461  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4311

Query Match 3.4%; Score 44; DB 4; Length 1461;  
Best Local Similarity 57.1%; Pred. No. 0.00038;  
Matches 80; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 909 TAGTTGTGTGGCCACGTTGGTGACTGTGTTGATCGCAGTTCAAACGCGATTGGCA 968  
DB 1405 TGTGTTCTGCTCGATGCTCGCGCACCTTACCGAAAGCATGTTCAAGCGCAATCCGAA 1346  
QY 969 TCAAGCATATGTCGAACCTTCTTCAGGCCACGGCGGATTCATGACCGTTTGGATGGCA 1028  
DB 1345 TCAAGCAGCAGCAACCTGCTGCCGCCACGGTGGCGTCTGATCGATCGACAGCC 1286  
QY 1029 TGCTCCCGCGCGATGCTG 1048  
DB 1285 TGACGGCAGCCATCCCGGTG 1266

RESULT 14  
US-09-543-681A-2949  
Sequence 2949, Application US/09/543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543.681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 2949  
LENGTH: 948  
TYPE: DNA  
ORGANISM: Proteus mirabilis  
US-09-543-681A-2949

Query Match 3.4%; Score 43.8; DB 4; Length 948;  
Best Local Similarity 46.3%; Pred. No. 0.00034;  
Matches 181; Conservative 0; Mismatches 207; Indels 3; Gaps 1;



